Functional dissection of the HNRPA2B1-CBX3 ubiquitous chromatin opening element (A2UCOE)

Anakok, Omer Faruk

Awarding institution:
King's College London

Functional dissection of the *HNRPA2B1-CBX3* ubiquitous chromatin opening element (A2UCOE)

by

Omer Faruk Anakok

Thesis submitted for the degree of Doctor of Philosophy

King's College London

Gene Expression and Therapy Group

Department of Medical and Molecular Genetics

Division of Genetics and Molecular Medicine

Faculty of Life Sciences and Medicine

King's College London

DECLARATION

I hereby confirm that, unless otherwise stated, the work presented in this thesis is my own.

Omer Faruk Anakok

Acknowledgements

I am indebted to many people for their support during this exciting, wonderful and mostly overwhelming process of working on my Ph.D.

First and foremost I am sincerely grateful to my advisor Dr Michael Antoniou. It has been an honour and privilege to work closely with him. I have enjoyed the opportunity to watch and learn from his knowledge and experience. His endless support, guidance, encouragement and patience with me are always appreciated. Millions of thanks would not even start to cover my gratitude to him.

Secondly I would like to thank my second supervisor Professor Gillian Bates for her support and constructive criticism during thesis compelling process.

I also would like to acknowledge the research committee members Dr. Steven Best, Dr. John Maher and Dr. Deborah Cunninghame Graham for their valuable feedback and contribution.

In addition, I would like to acknowledge Professor Agi Grigoriadis and Miss Ewa Kania from the department of Craniofacial development and stem cell biology, Kings college for their assistance and collaboration.

Dr. Fang Zhang and Prof Adrian trasher from molecular immunology unit, Institute of Child health, UCL

I consider myself very lucky to have cross paths with my laboratory colleagues: Christina Flouri, Robin Mesnage, Alexia Phedonos. Their help and friendship are much appreciated.

I am also grateful to the students I had the pleasure to work with during the time course, Kristian Skipper, Gautom Baruah.

I gratefully acknowledge the funding received towards my PhD from the ministry of education of Turkey, and also President Recep Tayyip Erdogan who made this sponsorship program possible.

Last but not least I would like to thank my long suffering other half Beyza Bedir who kept me sane for the last 4 years. Thanks for being my muse, editor, proof-reader and best friend during the darkest times of the whole PhD period. This PhD is a testament to her faith in me.

Abstract

Background

The ubiquitous chromatin opening element (UCOE) derived from the human *HNRPA2B1-CBX3* housekeeping gene locus (A2UCOE) is able to provide highly reproducible and stable expression even from transgenes integrated within extreme heterochromatic regions (such as centromeres). The A2UCOE consists of the methylation-free CpG island encompassing the dual divergently transcribed promoters of *HNRPA2B1* and *CBX3*. The proposed mechanism of A2UCOE function is a two-component model: (i) an extended methylation-free CpG island and (ii) dual divergent transcription with an inherent chromatin opening capability from the innate *HNRPA2B1* and *CBX3* promoters. Stable expression by the A2UCOE can be achieved from either driving transcription directly off the *HNRPA2B1* promoter or by linkage to a silencing-prone heterologous ubiquitous or tissue specific promoter.

The A2UCOE has been shown to provide highly reproducible and stable transgene expression from within lentiviral vectors (LVs) both *in vitro* and more importantly *in vivo* following *ex vivo* gene transfer to mouse bone marrow haematopoietic stem cells (HSCs). It has recently been shown that efficient and stable expression of transgenes from the A2UCOE is at least in part due to its resistance to DNA methylation-mediated silencing.

Project Aims

The aim of this project was two-fold. First, to test the two-component model of A2UCOE function by evaluating other elements with a similar CpG island - dual divergent transcriptional promoter configuration. Second, to dissect the A2UCOE in an effort to identify sub-regions that are crucial for its dominant chromatin opening activity and thus obtain a minimal fully functional element. This would allow further savings of space for therapeutic gene sequences once incorporated within LVs.

Methodology

Novel candidate UCOEs and A2UCOE sub-fragment-based eGFP reporter gene constructs within an LV context are functionally tested by transduction of P19 and F9 murine embryonic carcinoma cell lines and measurement of stability of expression over time both before and after differentiation down neuroectodermal and endodermal lineages respectively. Candidates showing encouraging signs of stability of expression in P19 and F9 cells would then also be tested for stability of expression in murine embryonic stem cells.

Results and Conclusions

The methylation-free CpG island, dual divergently transcribed regions used to test the overall two-component model of UCOE function were the native genomic *SETD3-CCNK* housekeeping gene pair and the artificially constructed single promoter *RPS11-HNRPA2B1* combined in a divergent configuration. Linkage of these elements in both orientations upstream of the highly silencing prone SFFV-eGFP reporter gene

system resulted in significant but only partial protection against silencing compared to the fully functional core 1.5A2UCOE-SFFV-eGFP reference. This was the case in P19 and F9 cell systems both before and after differentiation.

Various sub-fragments of the A2UCOE region ranging in size from 450-950bp either with or without associated promoter activity, were linked upstream of the SFFV-eGFP cassette and again compared to 1.5A2UCOE-SFFV-eGFP to assess their ability to negate transcriptional silencing. The results showed that none of the A2UCOE sub-fragments tested retained a full UCOE capability. Fragments from the *CBX3* first intron lacking promoter activity, but with a high CpG content, were particularly noteworthy for their total inability to rescue expression from the silencing-prone SFFV promoter. A 0.9kb sub-core fragment of the 1.5kb A2UCOE extending over both *CBX3* and *HNRPA2B1* transcriptional start sites was only partially able to negate silencing of the linked SFFV promoter.

Analysis of a deletion series from the *CBX3* end of a fully functional 2.2kb A2UCOE where expression of an eGFP reporter gene is directly driven off the *HNRPA2B1* promoter, revealed a 1.7kb truncation that retained full UCOE activity following transduction of P19 and F9 cells in both undifferentiated and differentiated states. A further deletion to 1.2kb lacked this capability. This 1.7A2UCOE was also able to retain stable expression in murine embryonic stem cells and during differentiation into embryoid bodies.

In summary, we were able to identify a native (*SETD3-CCNK*) and construct an artificial (*RPS11-HNRPA2B1*) dual divergent promoter combination with a significant but partial UCOE function compared to the fully functional prototypical *HNRPA2B1-CBX3* system. A2UCOE sub-fragments that were CpG rich but devoid of one or both of the *CBX3* and *HNRPA2B1* promoters, were at best only partially capable of conferring stability of expression. Lastly, we were able to substantially reduce the length of the A2UCOE required for full activity when gene expression is sought directly from the potent innate *HNRPA2B1* promoter.

Contents

Abbreviations

α	alpha
^{0}C	degrees Celsius
β	beta
Amp	ampicillin
bp	base pair
BSA	bovine serum albumin
cDNA	copy deoxyribonucleic acid
cPPT	central polypurine tract
CpG	cytosine/guanine DNA doublet
dH$_2$O	distilled water
DMEM	Dubecco's modified eagle medium
DNA	deoxyribonucleic acid
dNTPs	dinucleotide triphosphate
dsDNA	double-stranded deoxyribonucleic acid
eGFP	enhanced green fluorescent protein
EDTA	ethylenediaminetetraacetic acid
Env	envelope
FACS	fluorescent analysis cell sorting
FCS	foetal calf serum
Γ	gamma
g	gram
GAPDH	glyceraldehde-3-phosphate dehydrogenase
h	hour
HSC	hematopoietic stem cell
HIV-1	human immunodeficiency virus- type 1
ID	integration deficient
In	integrase

kb	kilo base pairs
LB	Luria Bertani
LCR	locus control region
LTR	long terminal repeat
LV	lentiviral vector
LVs	lentiviral vectors
µg	microgram
µl	microlitre
M	molar
mg	milligram
min	minutes
ml	millilitre
MA	matrix
MDR	Methylation determining region
MFI	mean fluorescence intensity
MOI	multiplicity of infection
mRNA	messager RNA
OD	optical density
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PIC	pre-integration complex
PPT	polypurine tract
qPCR	quantitative polymerase chain reaction
RNA	ribonucleic acid
RRE	*rev* responsive element
rpm	revolutions per minute
RT	reverse transcriptase
SIN	self-inactivating lentiviral vector
SFFV	spleen forming focus virus

Ss	single stranded
TAE	tris-acetate-EDTA
TU/ml	transfection units per ml
UCOE	ubiquitously-acting open chromatin element
v/v	volume per volume
WPRE	woodchuck post-transcriptional regulatory element

Chapter 1

Introduction

Chapter 1

Introduction

1.1.Gene Therapy

Gene therapy can be defined as the treatment of a disease or medical disorder by introducing of therapeutic nucleic acid material (DNA, RNA) into living cells. This nucleic acid material or therapeutic genes can change a specific gene mutation or re-program cell functions to treat a disease. Gene therapy can be achieved by viral and non-viral delivery (vector) methods. Viral vectors take advantage of the natural infectious properties of viruses whilst non-viral methods use DNA-pharmaceutical formulations such as liposomes.

1.1.1. Non-viral gene therapy

Non-viral gene therapy is the concept of introducing naked DNA, RNA and oligonucleotides into the cell. Although there are certain advantages of this methodology compared with viral methods, such as low host immunogenicity, and simple large-scale production, non-viral gene therapy has actually failed in most instances to present sufficient delivery and stable expression to be competitive to viral delivery. The one condition where non-viral vectors have shown some promise is targeting the lung in cystic fibrosis patients (Griesenbach and Alton 2013).

1.1.2. Viral gene therapy

Viruses are intracellular obligate parasites that evolved as efficient vehicles for the transfer of DNA or RNA by infection of target cells. Wild-type viruses infect the target cells by introducing their genetic material into the host cell as part of their replication cycle. The genetic material of wild-type viruses, takes control the target cell normal replication mechanisms to serve the requirements of the virus and the target cells in the end follow these orders and produce additional copies of virus, causing more and more cells to become infected. This process of infection opened the way for viruses to be used as vectors to carry desired genetic material into human cells. Additionally, the removal of genetic disease-causing viral genes from the vectors and the replacement of these parts with genes encoding therapeutic genes of interest showed a viable and safe alternative method to common therapeutics, and provided a way forward in the field of gene therapy.

Today, a large number of viruses with exclusive characters valuable for gene therapy have been identified. This led to the use of recombinant viruses such as adenoviruses, adeno-associated viruses, herpes viruses, poxviruses, retroviruses and more recently lentiviruses both in the laboratory and clinic.

1.1.2.1.Adenovirus

Adenoviruses are non-enveloped viruses having a linear double stranded DNA genome. Today, over 40 strains of adenovirus have been identified, and most of them are known as causing benign respiratory tract infections in humans. Adenovirus genetic material normally replicates as an episomal element in host cells instead of integration into the host genome. The adenovirus genome is approximately 35kb, which in the case of "gutted" vectors can be replaced with approximately 30kb foreign DNA (Verma and Somia 1997).

Although, adenoviruses are very efficient at transducing a wide range of target cells *in vitro* and *in vivo*, and can be produced at high vector titres ($>10^{11}$/ml), transgene expression *in vivo* can be transitory due to the fact that this class of virus can elicit a potent immune response, which can limit their therapeutic application (Verma and Somia 1997).

1.1.2.2.Adeno-associated virus

Adeno-associated viruses (AAV) are a non-pathogenic group of human viruses that are normally dependent on a helper adenovirus to propagate. They have a single stranded DNA genome, consisting of two genes; *rep*, coding of for proteins that control viral replication, structural gene expression and integration into the host genome; and *cap*, which codes for capsid structural proteins.

As in the case of adenoviral vectors, AAV gene therapy vectors have the ability to infect both dividing and non-dividing cells. They are retained in an extra chromosomal state without recombining into the genome of the host cell. Some studies have reported a degree of clinical success with AAV vectors targeting the retina in cases of Leber's Congenital Amaurosis (Bainbridge, Tan et al. 2006, Alexander, Linde-Zwirble et al. 2007).

1.1.2.3.Retrovirus

Retroviruses are viruses characterised by three main features: they have a single stranded (ss) RNA genome; possess their own reverse transcriptase enzyme; and a virion morphology that consists of two proteinaceous structures and a dense core, all surrounded by an envelope structure (Vogt and Simon 1999).

Retroviruses have a lipid bilayer-enveloped particle structure and their ssRNA genome can vary from 7 to 11kb in length. Upon cell entry, retroviruses move from the cell surface to the nucleus. In the mean time the ssRNA genome is retro-transcribed into linear double stranded (ds) DNA by their reverse transcriptase enzyme. When dsDNA reaches the nucleus, it integrates into the host cell genome to form a structure known as the "provirus" and is passed on to daughter cells upon cell division. The proviral genome consists of long terminal repeats (LTRs) located at both ends (5' and 3') containing various components such as the U3 and U5 transcriptional control elements including promoters and enhancers (Figure 1.1). The viral LTRs flank the viral genome

with the 5'LTR responsible for transcription of the tandem genes encoding group-specific antigen (gag) that form the core and structural proteins of the matrix and nucleocapsid, polymerase (pol) coding for the reverse transcriptase protease and integrase enzymes, and envelope (env), which is a transmembrane protein component of the retroviral coat. The other parts of the retroviral genome are the packaging signal (psi or ψ) required for the specific packaging of the full length viral RNA genome into newly formed virions (Watanabe and Temin 1982), and the central polypurine taract (cPPT), which is the site of the initiation of positive-strand DNA synthesis during reverse transcription (Rattray and Champoux 1989, Charneau, Alizon et al. 1992).

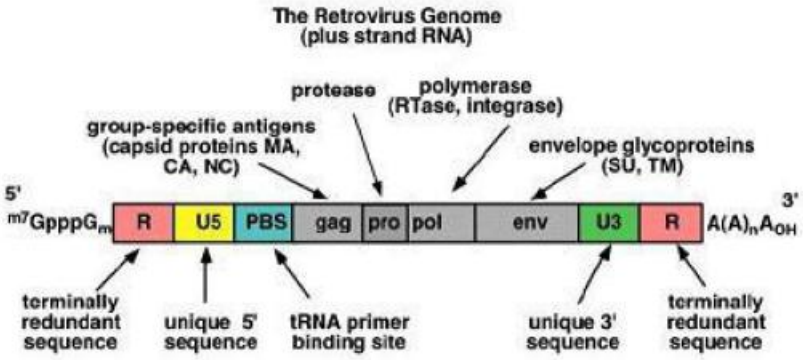


Figure 1.1. General structure of a retroviral genome. The long terminal repeats (LTRs) have sequences needed for the regulation and initiation of transcription within the unique sequence 3 and 5 (U3, U5) regions. (Image taken from *http://envmedical.com*)

The *Retroviridae* are divided in two family groups namely simple and complex retroviruses. MLV oncoretroviruses also known as gammaretroviruses are examples of the simple class and lentiviruses such as human immunodeficiency virus type-1(HIV)-1 typifies the complex group of retroviruses.

1.1.2.4.Lentiviruses

Lentiviral vectors (LVs), are most frequently based on HIV-1. LVs possess advantages over other types of viral vectors based on gammaretroviruses as they are capable of transducing non-dividing as well as dividing cells (Naldini, Blömer et al. 1996). The property of LVs to transduce non-dividing cells is clearly a great asset when targeting either post-mitotic cells or tissues such as those of the central nervous system and liver or slow mitotic stem cell

populations. In addition, the capability of retroviral/LVs to integrate their genetic material into the target cell genome is a clear advantage when targeting cells that undergo division, especially stem cells, as this at least decreases and possibly avoids the need for their repeated administration.

Another advantage of using LVs is that they are relatively large with a manageable capacity of approximately 8kb. Furthermore, by pseudotyping LVs with envelope glycoproteins of other viruses during vector production, it is possible to redirect their tropism to a broader range of cell types (Bouard, Alazard-Dany et al. 2009). In addition, the use of cell- and tissue-specific gene regulatory elements (LCRs, enhancers, promoters) (Lai and Brady 2002, Greenberg, Geller et al. 2007, Gascón, Paez-Gomez et al. 2008) can provide a restrictive transgene expression pattern and thus reduce the risk of immune reactions against the therapeutic product.

There have been enormous improvements in the safety of LVs since they were first developed. The wild type viral genes of the first generation (Naldini, Blömer et al. 1996) (VSV 1996) and non-essential parts of the second and third generation genome have been removed (Loeb, Cordier et al. 1999). This caused a lack of a replicative capacity, which has overcome the problem of producing infectious particles. Viral production is achieved with a packaging cell line, which is transfected with three or four plasmids, across which the sequences to make a vector are distributed. In this production system, only a small part of the original vector genome remains even across this allocation of components

between different plasmids and as such it prevents the possible homologous recombination between the plasmids to reconstruct replication-competent virions. This is in addition to avoiding the risk of the possible rescuing of the vector in HIV-infected individuals. Only the transfer vector sequences, which is one of the 3 or 4 transfected plasmids and contains the therapeutic transgene cassette, is packaged into virions along with the required proteins for reverse transcription and integration of the vector genome into the target cell. The remaining viral genes whose products are required for replication lie on the helper plasmids that remain within the packaging cell line.

1.1.3. HIV virus

1.1.3.1.HIV genome structure

The wild type HIV genome compromise of two unspliced ssRNA molecules of approximately 10 kb length located within a nuclear core enclosed by a nucleocapsid and a viral envelope core.

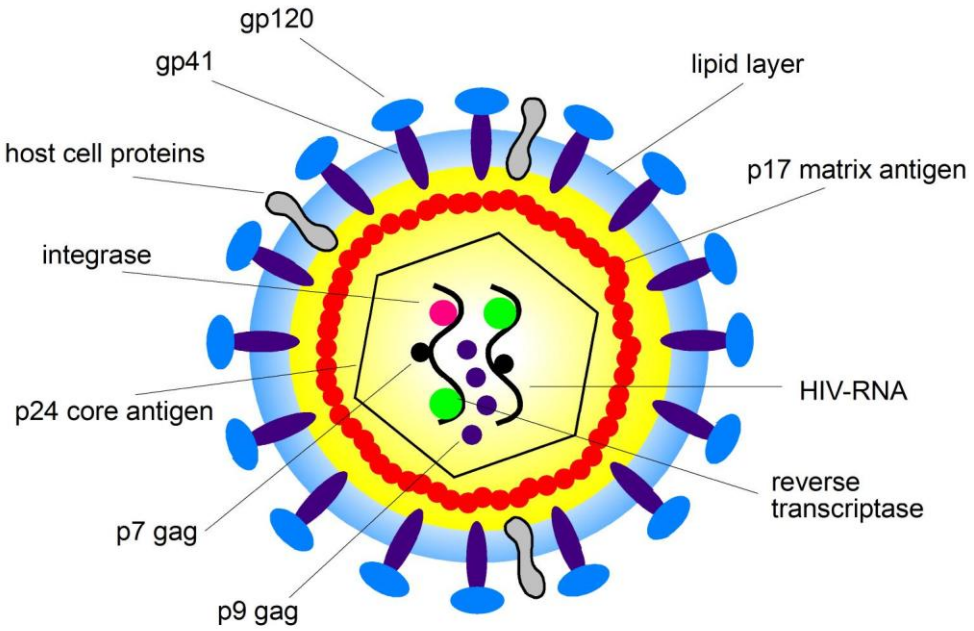
The HIV genome is regularized into the *gag*, *pol*, and *env* gene form (Figure 1.2). Gag gene roles on encoding the structural proteins, and the pol gene provides the enzymes that comes with the ssRNA. The reverse transcriptase enzyme performs the reverse transcription of the viral RNA to DNA. The integrase enzyme carries out the integration of the proviral DNA into the host genome, and protease enzyme regulates the *gag/pol* disintegration and virion maturation (Katz and Skalka 1994). Env roles on encoding the viral envelope

part of the virion.

The Wild type HIV-1 has two regulatory genes, *tat* and *rev*. These genes are necessary for viral replication and gene expression. The protein *tat* gene encoding plays in the activation of the promoter in the HIV LTR so that viral RNA is produced efficiently, and Rev interacts with a region of viral RNA known as the Rev-responsive element (*rre*) and promotes the transport of viral RNA from the nucleus to the cytoplasm (Pluta and Kacprzak 2009). Although HIV-1 holds the retroviral genome organisation, it also has four additional accessory genes, *vif, vpr, vpu* and *nef,* which play for *in vivo* replication and pathogenesis (Figure 1.2; (Coil and Miller 2004)).

The later researches have focused on the use of the therapeutic vectors raises safety concerns as the cytotoxic or cytostatic encoding proteins; *vpr* induces G2 cell cycle and *nef* alters cellular activation pathways. Cell surface molecules such as CD4 and the class I major histocompatibility complex are induced by *nef* and *vpu*. In addition, the accessory genes *nef*, *vif* and *vpr* are integrated into viral particles and can improve the immunogenicity of vectors. For those safety concerns, lentiviral vectors (LV) are constructed in the absence of these genes, which have been shown not to be essential for viral transduction (Delenda 2004, Pluta and Kacprzak 2009).

A)



B)

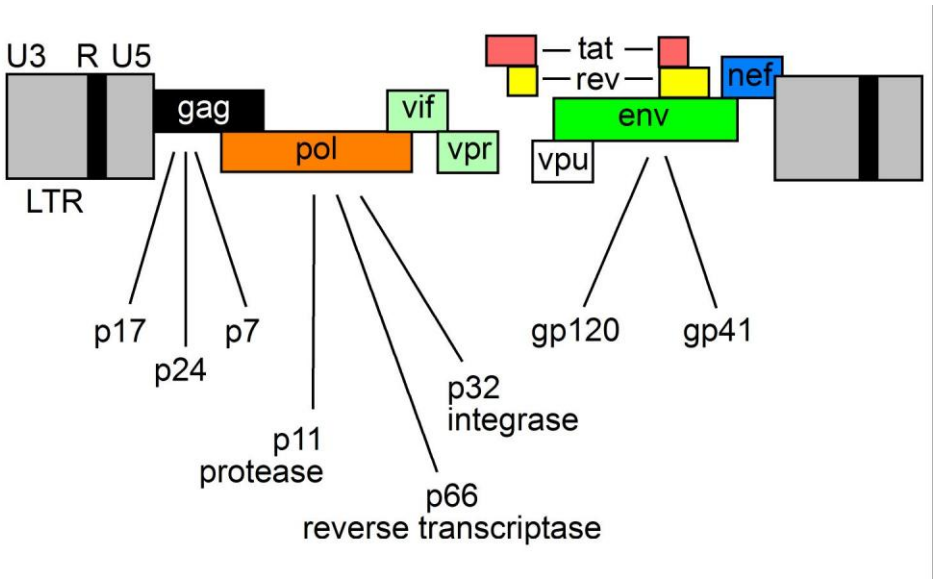


Figure 1.2. The wild type HIV-1 genome. A) demonstarates the 2D representation of the virus, with the encoding proteins (gp41, p17, p24, p7.,etc.) encircling te core containing the viral RNA and reverse transcriptase. The virus is mainly composed by three genes (*Gag*, *Pol* and *Env*) encoding the packaging and envelope proteins of the virus flanked by the two LTRs, as shown in **B)** (Image from *http://hivbook.com/tag/viral-genome/*)

1.1.3.2. HIV life cycle

The HIV life cycle starts by binding of the viral envelope glycoprotein gp120 to a CD4 receptor and a secondary receptor on the cell surface after which conformational changes in the non-covalently associated gp41 subunit releases free energy sufficient to endorse fusion of the viral particle envelope with the cell membrane. The viral core is then transferred into the cell cytoplasm and reverse transcription of the viral RNA takes place. The reverse transcription unit (RTC) is then moved towards the nucleus by the cellular proteins. When reverse transcription cycle is finished and the provirus and several viral proteins formed, the pre-integration complex (PIC) is then transferred into the nucleus by the cellular proteins. After the integration to the cell genome, transcription is carried out by RNAP II and enhanced by the viral protein Tat. Unspliced and partially spliced transcripts are stabilised by the viral protein *Rev* (Pluta and Kacprzak 2009).

1.1.4. Development of lentiviral vectors

For safety concerns, reducing the risk associated with producing and use of LVs in gene therapy, all non-essential genes coding for accessory proteins and responsible for virulence, should be removed from the vector sequence. Moreover, the vector genome is typically broken up into several parts with limited sequence overlap to minimize the possibility of independent recombination to form wild type viral particles and vector mobilization. The

genes necessary for lentiviral vector production are therefore divided on different plasmids that are co-transfected into HEK293T cells together as a complex aggregated with the help of a transfection agent (see Materials and Methods section). Therefore, the generation system typically consists of packaging expression cassettes (helper), a vector cassette (transfer vector) and an envelope expression cassette (Figure 1.3).

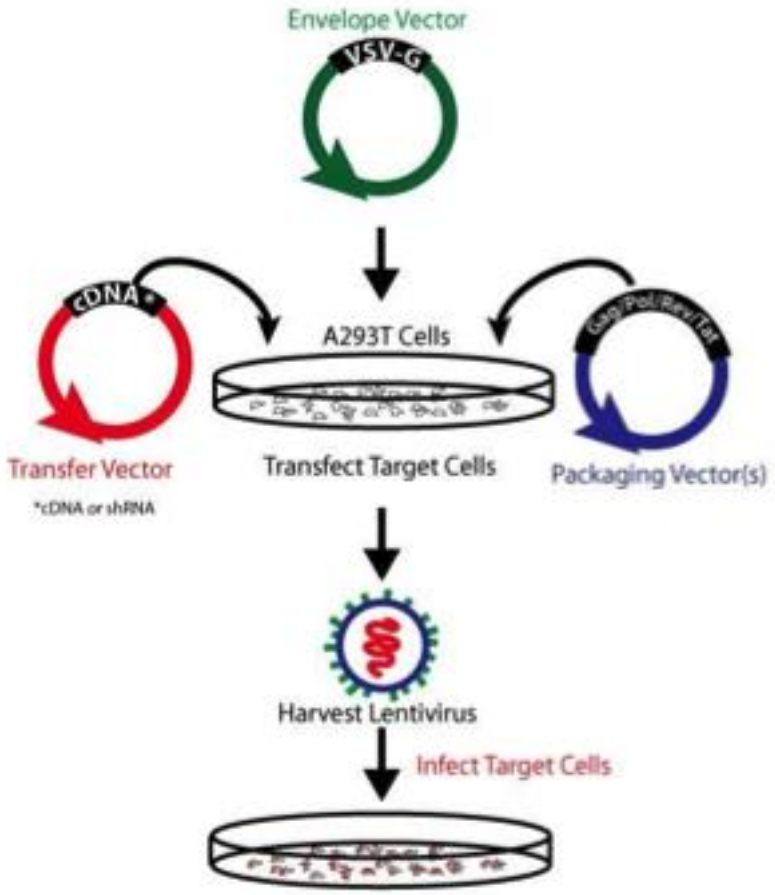


Figure 1.3. Lentiviral vector production protocol by 293T cells. As in oncoretroviral vectors, the viral genes *gag*, *pol* and *env* are replaced by the transgene of interest. The LTR and the packaging signal ψ are maintained for vector amplification and packaging. Of the lentiviral accessory proteins, only the *rev* responsive element (RRE) is maintained as an additional *cis*-acting sequence required for the nuclear export of unspliced and single-spliced viral RNAs in the presence of the Rev protein. Rev is expressed by one of the three

packaging constructs. The two others encode the Gag/Pol or the Env proteins, respectively. (Image taken from http://www.addgene.org/)

1.1.5. Development of SIN vectors

The risks associated with viral recombination can happen at the DNA level, particularly during the production of co-transfected plasmids, between transfected plasmids or between the proviral vector and a homologous chromosomal sequence in the target cell. Therefore, the possibility of full-length LV transcription and propagation after integration into the target cell needs to be addressed. It is possible for the integrated proviral LV to be mobilized by replication competent virus; such as, if the transduced cell were later infected with wild type HIV-1. In addition, there is also the possibility of the unusual expression of host genes by LTR that are adjacent to the site of LV integration. These concerns led to the development self-inactivating (SIN) LVs.

The first SIN retroviral vectors were prepared in an MLV- based vector (Yu, Nash et al. 1986, Yee, Pasi et al. 1999). The first vector was produced with a deletion of 299bp in the U3 region, containing the enhancer and promoter CAAT box unit with the second having the additional deletion of the TATA box. Nevertheless, retroviral vectors exhibit weak polyadenylation sites and deletion of U3 to generate SIN vectors increased the chances of read-through thus augmenting the potential for insertional (Furger, Monks et al. 2001, Zaiss, Son et al. 2002). The same method was used later in LVs design by *Miyoshi* and colleagues (Miyoshi, Blömer et al. 1998) by further deletion of a 133bp section from the U3 region of the 3‟ LTR, which removed the TATA box and other transcription factor binding sites, resulting in transcriptional inactivation of the

proviral LTR in targeted cells, both *in vitro* and *in vivo.* This modification demonstrated no changes in the viral transcript levels of the producer cells, and no major reduction in viral titre. In addition, expression of the transgene *in vivo* in both neurons and retinal cells was improved with this SIN vector due to the removal of transcriptional interference from the HIV-1 LTR promoter/enhancer by this research. The study also reported a further deletion of the enhancer/promoter sequences in the U3 region of the 3‟ LTR (120-40bp U3).

Additional success was obtained by Zufferey and colleagues (Zufferey, Donello et al. 1999) with a deletion of up to 400bp of the 3'LTR U3 region. Once more, the vector production was neither decreased nor was transduction efficiency lesser *in vitro* or *in vivo*. In addition, it has been shown that in SIN vector transduced cells that were subsequently infected with wild type HIV-I, the vector was not mobilised (Poccia, Battistini et al. 1999).

By the development of SIN LV, promoter interference was extensively reduced and this also provided a potential safety advantage over traditional retroviral vectors. In particular, upon reverse transcription, during the step of viral replication, the U3 region will be duplicated in both LTRs, thus both LTRs are inactivated leaving only the very 5‟ end of the U3 for integrase detection and function. Expression of the transgene will thus be restricted from an internal promoter.

Transcription of full-length viral vector RNA transcripts is noticeably reduced in cells transduced with a SIN vector when the LTR deletion is sufficient. This

also decreases the risk of generating replication competent retrovirus and there is a rather incomplete U3 enhancer sequence in the proviral 5'LTR interfering with internal, heterologous promoters. Moreover, it provides additional safety advantages, as it is more difficult to re-generate a wild-type parental retrovirus via recombination. With the deletion of the viral transcriptional elements from the vector, synthesis of vector RNA will depend on site of integration (Kappes and Wu 2001).

Additionally, the SIN vector design avoids transcriptional interference by the promoter/enhancer units in the host genome, reducing the possibility of insertional activation and mutagenesis of adjacent coding sequences (such as oncogenes) either by transcription from the 3'LTR at the site of vector integration or LTR enhancer activation of a host gene promoter. To improve biosafety, the use of SIN vectors also provides two additional advantages; elimination of transcriptional interference by the LTR promoter, and the possibility to create tissue-specific and inducible vectors via appropriate internal promoters, which would be not easy in the presence of non-specific transcription from the LTR promoter (Kappes and Wu 2001).

1.2.Success in clinical trials using viral vectors

1.2.1. Trials of gene therapy for X-linked severe combined immunodeficiency (SCID-X1)

X-linked severe combined immunodeficiency disease (SCID-X1) is an inherited disease with absent T lymphocyte function, due to lack of the common cytokine

receptor γ-chain (γc; IL2RG). In 2000, a successful treatment of 2 children suffering from SCID-X1 was achieved by the transduction of a normal functioning copy of the IL2RG cDNA in patient bone marrow HSC using a gammaretroviral vector based on murine Maloney leukaemia virus by an *ex vivo* delivery method (Cavazzana-Calvo, Hacein-Bey et al. 2000, Hacein-Bey-Abina, Le Deist et al. 2002). A similar and equally successful gene therapy approach for SCID-X1 was also carried out in London (Gaspar, Parsley et al. 2004). During the years following treatment, results demonstrated no indication of replication competent retrovirus in the patients. All patients responded well to the treatment with restoration of immune system function. However, of the 20 patients treated, 5 subsequently developed a lymphoproliferative, leukaemia-type condition associated with the gene therapy itself. The analysis of these cases found that the root cause was gammaretrovirus-induced insertional mutagenesis leading to upregulation of proto-oncogene expression, particularly *LMO2* (Hacein-Bey-Abina, Von Kalle et al. 2003, Howe, Mansour et al. 2008, Escors and Breckpot 2010).

1.2.2. Gene therapy for chronic granulomatous disease (CGD)

A treatment for X-linked chronic granulomatosis disease (X-CGD) was reported in 2006, using a gammaretroviral vector carrying the therapeutic gp91phox cDNA under control of the SFFV promoter-enhancer element within the vector LTR and targeting bone marrow HSC (Ott, Schmidt et al. 2006). A successful outcome was initially observed in both adult male patients treated in this trial using an *ex vivo* procedure. The two patients used in this treatment showed a

high level of corrected leukocytes in peripheral blood (~20%) within the first months after treatment, which reached up to 80% over the first year. However, a benign form of myelodysplasia, a pre-leukemic condition was then seen in the patients by the insertion within and subsequent activation of the MDS1-EVI1, PRDM16 and SETBP1 genes (Stein, Ott et al. 2010) via the vector SFFV enhancer. Crucially, although neutrophil counts were observed to be maintained at normal levels, therapeutic transgene function was lost in both cases leading to death of one of the patients from severe bacterial sepsis after colon perforation 27 months after gene therapy. Molecular analysis showed that in both patients the SFFV promoter (but not enhancer) driving expression of the therapeutic gp91phox had become methylated and thus inactivated leading ultimately to treatment failure (Ott, Schmidt et al. 2006, Stein, Ott et al. 2010). A possible solution to this problem of transgene silencing and treatment failure is to employ therapeutic gene cassettes whose expression is augmented with the ubiquitous chromatin opening element (UCOE) from the human *HNRPA2B1-CBX3* locus (Brendel, Müller-Kuller et al. 2011).

1.2.3. Gene therapy for Leber's Congenital Amaurosis (LCA)

A number of clinical trials with encouraging outcomes have targeted the genetic eye disorder Leber's Congenital Amaurosis (LCA). Around 10% of LCA patients are defective in *RPE65*, whose protein product is required for normal retinal cycling of vitamin A. Thus in the reported trials AAV vectors were employed to deliver *in vivo* a correctly functioning copy of the RPE65 cDNA to the retinal pigment epithelia cells (Bainbridge, Tan et al. 2006, Alexander,

Linde-Zwirble et al. 2007). The results from three young adult patients with severe retinal dystrophy resulted in one showing a marked improvement in retinal function and visual mobility in low light conditions (Bainbridge, Tan et al. 2006).

In a second gene therapy trial for LCA using an AAV-RPE65 vector also showed signs of improvement in visual perception in all three patients treated (Maguire, Simonelli et al. 2008). In addition, the safety and efficacy of the vector has continued through to at least 1.5 years post-injection (Simonelli, Maguire et al. 2009).

1.2.4. First human gene therapy trials using lentiviral vectors – gene therapy for X-linked adrenoleukodystrophy (X-ALD) and metachromatic leukodystrophy (MLD)

The first clinical therapy using a lentiviral vector for an inherited disease targeted X-linked adrenoleukodystrophy (X-ALD) via an *ex vivo* HSC approach (Cartier, Hacein-Bey-Abina et al. 2009). X-ALD is a brain demyelinating disease caused by mutation of *ABCD1* encoding the ALD protein that is an adenosine triphosphate binding cassette transporter in the membrane of peroxisomes. The LV used in this treatment expressed a wild-type ABCD1 cDNA under the control of the myeloproliferatuve sarcoma virus promoter-enhancer (MDN) and was used to genetically correct autologous cytokine-mobilised peripheral blood $CD34^+$ cells *ex vivo*. In this first trial two young patients were transplanted with these autologous corrected HSCs following pre-

transplant conditioning with myeloablative doses of busulfan and cyclophosphamide. The results showed that this therapeutic approach stopped progression of the disease with the same efficacy as transplant of allogeneic bone marrow HSC from a matched donor. Thus, this study demonstrated that LVs could mediate important improvements of gene transfer into human HSCs.

Following on from this success with X-ALD, a similar procedure was employed to treat patients with metachromatic leukodystrophy (MLD; (Biffi, Montini et al. 2013)). MLD is an autosomal recessive inherited, lysosomal storage disease condition. It is caused by mutations in *ARSA* gene giving rise to deficiency of arylsulfatase A. This in turn results in accumulation of the enzyme substrate sulfatide within numerous cell types namely oligodendrocytes, microglia, and certain neurons of the central nervous system, and Schwann cells and macrophages of the peripheral nervous system. The build-up of sulfatide gives rise to widespread demyelination and neurodegeneration coupled with severe progressive motor and cognitive impairment. The condition is fatal within a few years of disease onset.

An LV containing an ARSA cDNA under control of the phosphoglycerate kinase promoter (PGK.ARSA.LV) was used in an *ex vivo* bone marrow HSC gene therapy procedure to treat a total of 9 patients (6 infants, 3 older) who also underwent full myeloablative conditioning with busulfan prior to autologous gene corrected HSC transplantation. Polyclonal engraftment in all peripheral blood lineages was obtained in all cases with stable *ARSA* enzyme activity

observed including within the CNS. Overall, a marked improvement in clinical condition was observed in all cases including decline in peripheral neuropathy (proportional to level of engraftment), total rescue of motor ability over time and normal cognitive ability for the age of the patient.

1.3.Gene silencing

Gene silencing is a generic term that includes a wide range of related topics. A frequent effect observed with embryonic carcinoma (and other) cells transduced with retroviral vectors is transcriptional silencing shortly after infection (Teich, Weiss et al. 1977, Teich, Weiss et al. 1977, Speers, Gautsch et al. 1980, Ellis 2005). Another problem that can also occur following retroviral vector transduction is variegation, in which genetically identical sister cells containing the same provirus may either express or be silenced (Yao, Sukonnik et al. 2004). So this situation refers to the progressive silencing of an initially expressed provirus during long-term culture or differentiation (Laker, Meyer et al. 1998).

1.4.Silencing of lentiviral vectors

LVs were first developed by Naldini *et al.* in 1996, with subsequent improvements in vector design and packaging systems (Zufferey, Dull et al. 1998). Subsequently there has been a large accumulation of data concerning gene silencing of self-inactivating (SIN) LVs. Transgene silencing can be through one of two mechanisms. The first is through integration within a region

of heterochromatin. Second, *de novo* methylation of the transgene promoter can take place despite the fact that gammaretroviral and lentiviral vectors preferentially integrate within active, transcriptionally permissive regions of the genome (Antoniou, Skipper et al. 2013).

For example, a SIN PGK-EGFP LV was shown to allow the non-selective isolation of mouse ES cells transduced at a single proviral copy (Yao, Sukonnik et al. 2004). The PGK-EGFP LV in the ES cell lines obtained was silent, variegated or low expressing, and in addition, subject to extinction during differentiation to embryoid bodies. Moreover, chromatin immunoprecipitation experiments showed that silent chromatin composed of bound histone H1 and deacetylated histone H3 preferentially forms on the HIV-1 LV sequences to silence the EGFP reporter gene, similar to the silent chromatin that forms on gammaretroviral vectors. These findings establish that SIN LVs can exhibit all aspects of gammaretroviral gene silencing in ES cells, and the same chromatin features are observed with silenced LVs (Ellis 2005).

1.5.Ubiquitous Chromatin Opening Elements (UCOEs)

When producing a therapeutic protein, a number of different factors affect the process. The level of RNA expression depends on the design of the vector used. Many mammalian expression vectors are commercially available, which incorporate different regulatory elements to maximize the expression of the therapeutic protein including promoters and enhancers, and post-transcriptional

elements such as the woodchuck hepatitis virus post-transcriptional regulatory element (WPRE). Moreover, getting reliable and stable transgene expression in mammalian cells is another major challenge. The problem of mammalian vector system inadequacy to express proteins in a stable manner can frequently result in silencing of the exogenous gene usually by methylation of CpG DNA sequences and histone deacetylation leading to chromatin condensation (Razin 1998, Fuks 2005). The discovery of ubiquitous chromatin opening elements (UCOEs) that are able to confer resistance to heterochromatin and DNA methylation-mediated silencing has led to the incorporation of these elements into expression vectors including LVs (Benton, Chen et al. 2002, Benton, Chen et al. 2002, Ellis 2005, Zhang, Thornhill et al. 2007, Zhang, Frost et al. 2010).

UCOEs are currently defined as methylation-free CpG island elements that extend over at least one and preferably two closely spaced, dual divergently transcribed promoters, which are derived from housekeeping gene loci (Antoniou, Harland et al. 2003, Antoniou, Harland et al. 2003, Williams, Mustoe et al. 2005). The prototypical UCOEs are those derived from the human *TBP-PSMB1* and *HNRPA2B1-CBX3* housekeeping gene loci with both consisting of a methylation-free CpG island spanning dual, divergently transcribed promoters (Figure 1.4; (Antoniou, Harland et al. 2003, Williams, Mustoe et al. 2005)).

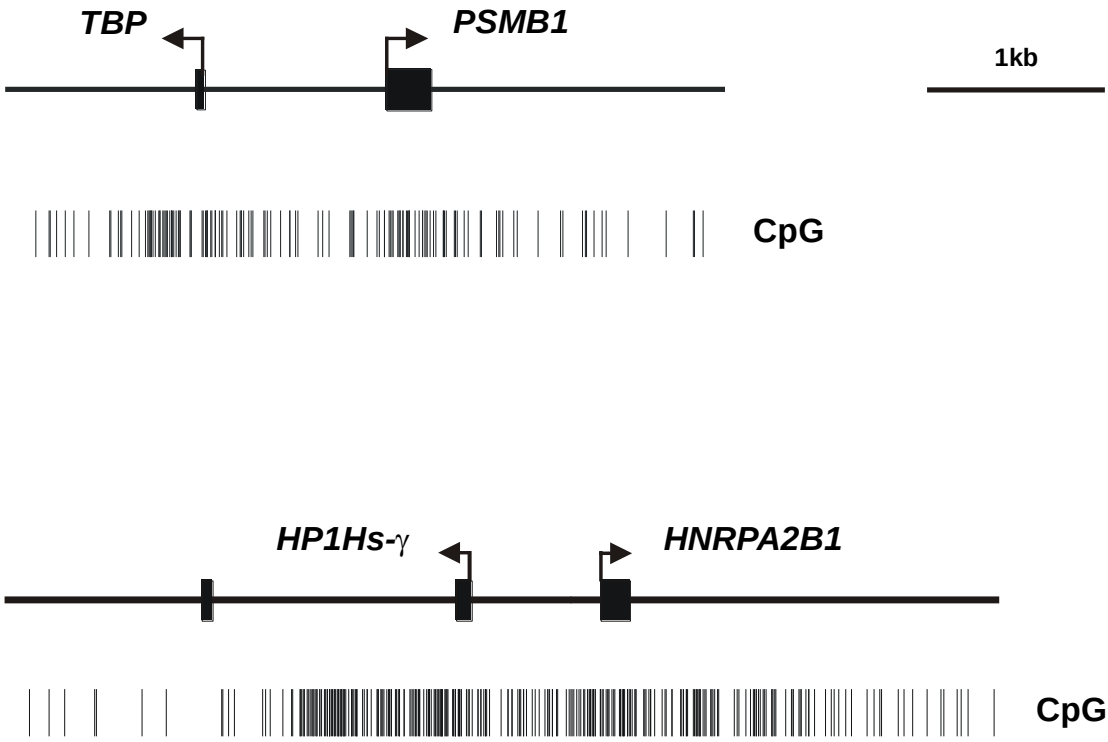


Figure 1.4. Illustration of the closely spaced, dual divergently transcribed promoter regions of the human *CBX3-HNRPA2B1* and *TBP-PSMB1* loci first identified to have UCOE function. Black boxes denote exons. Horizontal arrows show direction of transcription. CpG dinuclotide density maps are shown below the genomic structural illustration.

Using assays in stably transfected cell lines with plasmid vectors, both of these elements were shown to be capable of conferring reproducible and stable expression even from within centromeric transgene integration sites demonstrating their dominant chromatin opening capability (Antoniou, Harland et al. 2003). The UCOE from the *HNRPA2B1-CBX3* locus (A2UCOE; Figure 1.5) is able to provide stable transgene expression by transcription from the innate *HNRPA2B1* promoter (Antoniou, Harland et al. 2003). In addition, the A2UCOE can confer stable expression on linked silencing-prone promoters such as CMV (Williams, Mustoe et al. 2005). The proposed mechanism by

which the A2UCOE can provide stable transgene expression is a two-component model; an extended region of methylation-free DNA coupled with divergent transcription with inherent chromatin remodelling (Allen and Antoniou 2007).

In more recent years A2UCOE-regulated transgenes have been successfully integrated into LVs. As seen in stably transfected cell lines with plasmid vectors (Antoniou, Harland et al. 2003, Williams, Mustoe et al. 2005), the A2UCOE can provide stable expression either directly from the innate *HNRPA2B1* promoter (Zhang, Thornhill et al. 2007) or by linkage upstream of ubiquitous (Zhang, Thornhill et al. 2007, Pfaff, Lachmann et al. 2013) or tissue-specific (Brendel, Müller-Kuller et al. 2012) promoters. Stability of expression is observed not only in tissue culture but more importantly in HSCs and all their differentiated progeny *in vivo* (Zhang, Thornhill et al. 2007, Zhang, Frost et al. 2010, Brendel, Müller-Kuller et al. 2012). In addition, more recently A2UCOE-augmented cassettes in LVs have been shown to provide stable expression in embryonic and induced pluripotent stem cells and differentiated progeny representative of all three germ layers (Pfaff, Lachmann et al. 2013, Ackermann, Lachmann et al. 2014). A2UCOE resistance to silencing within LVs is at least in large part due to its ability to negate promoter DNA methylation (Zhang, Frost et al. 2010, Brendel, Müller-Kuller et al. 2012, Pfaff, Lachmann et al. 2013, Ackermann, Lachmann et al. 2014).

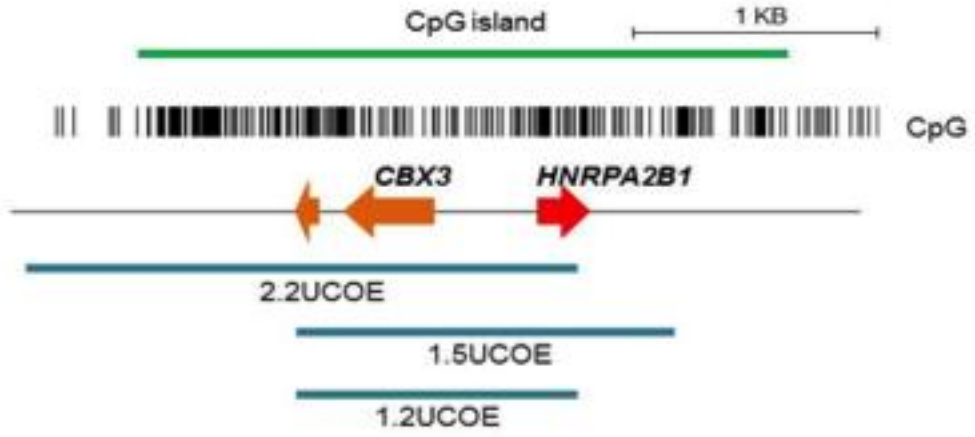


Figure 1.5. Illustration of the relative positions of the A2UCOE region sub-fragments shown to possess full UCOE activity. Upper panel: CpG density map and methylation-free CpG island. Middle panel: Positions of the two alternative first exons of *CBX3* (orange arrows) and the first exon of *HNRPA2B1* (red arrow). Lower panel: positions of the fully functional A2UCOE sub-region fragments. 2.2UCOE: 2.2kb element used to drive expression from *hnrpa2b1* promoter. 1.5UCOE: 1.5kb element used to link upstream of heterologous promoters. 12UCOE: a 1.2kb sub-fragment of the 1.5UCOE element with a deletion of 300bp from the *hnrpa2b1* end to match that in the 2.2UCOE fragment.

These properties of the A2UCOE suggest that it can potentially play a significant role in gene therapy LVs for a number of reasons. First, the A2UCOE provides stable transgene expression in the absence of an enhancer element, which reduces the possible risk of insertional mutagenesis and thus providing a higher safety profile. Second, A2UCOE-LVs are resistant to integration site position-specific reduction in expression of both the inherent A2UCOE and linked ubiquitous and tissue-specific heterologous promoters (Zhang, Thornhill et al. 2007, Zhang, Frost et al. 2010).

We and others have sought to identify any A2UCOE region sub-fragments that retain their resistance to transcriptional silencing. Prior to the commencement of this study, a 2.2kb A2UCOE fragment had been identified as the minimum required to drive stable expression from the innate HNRPA2B1 promoter (Antoniou, Harland et al. 2003). In addition, a core 1.5kb (Williams, Mustoe et al. 2005, Brendel, Müller-Kuller et al. 2012) or 1.2kb (Zhang, Frost et al. 2010) fragment from the A2UCOE spanning the transcriptional start sites of both *HNRPA2B1* and *CBX3*, were found to stabilise expression of linked heterologous promoters (Figure 1.5).

Recently, a 700bp sub-fragment of the 1.5kb A2UCOE extending over the *CBX3* promoter and alternative first exons alone has been claimed to confer at least partial resistance to silencing from both its innate albeit very weak *CBX3* promoter and when linked to heterologous *EF1α* and *MRP8* promoters (Brendel, Müller-Kuller et al. 2011). All these A2UCOE fragments function (Brendel, Müller-Kuller et al. 2011) through a mechanism that negates DNA methylation-mediation silencing.

Interestingly, both the 1.5kb and 1.2kb A2UCOE fragments (Figure 1.5) in an LV context targeting HSCs and embryonic stem cells were found to confer stability and specificity of expression on linked promoters in an orientation-dependent manner; that is, with the *CBX3* end of the element juxtaposed next to the downstream heterologous promoter (Zhang, Frost et al. 2010, Brendel, Müller-Kuller et al. 2011, Brendel, Müller-Kuller et al. 2011). Furthermore, it

has also been reported that a promoter-less sub-fragment of the A2UCOE derived from the region downstream of the CBX3 promoter within the first intron of this gene, is also capable of conferring stability on a linked heterologous promoter within an LV context (Bandaranayake, Correnti et al. 2011).

1.6.Project objectives

Given the current status of the field surrounding UCOEs, the overall objectives of this project were to functionally test current model of the UCOE mechanism of action, namely the requirement for dual divergent transcription from within an extended DNA methylation-free CpG island region. In order to address this we planned to:

1. Test the UCOE concept for its generality by identifying novel elements from housekeeping gene loci with a dual divergently transcribed promoter configuration but which would also function in an orientation independent manner.
2. Functionally dissect the A2UCOE region to identify sub-fragments that retain a full transcriptional stabilising function but which (i) potentially lack one or both of the HNRPA2B1 and CBX3 promoters and (ii) reduce the size of the associated CpG island.

All test constructs would be within an LV context and functional analysis by transduction and expression analysis within murine embryonic teratocarcinoma P19 and F9 cells as they are a proven model for gene silencing (Zhang, Frost et

al. 2010). Studies would be extended to those previously conducted by investigating stability of expression in not only undifferentiated cells but also in P19 cells induced to undergo differentiation into neurons and F9 cells that had undergone endodermal differentiation (Jones-Villeneuve, McBURNEY et al. 1982, Yoon, Lee et al. 2009). Finally, novel candidate UCOE constructs that gave encouraging results in P19 and F9 cell assays would be tested within a more physiological murine embryonic stem (ES) cell context.

Chapter 2

Materials and Methods

Chapter 2

Materials and Methods

2.1.Materials

2.1.1. Commercial kits

EndoFree Plasmid Maxi kit (QIAGEN®, *Hilden, Germany*)

HighSpeed Plasmid Maxi Kit (QIAGEN®)

QIAprep Spin Miniprep kit (QIAGEN®)

QIAzol Lysis reagent (QIAGEN®)

Wizard® Genomic DNA purification kit (Promega UK, Southampton, UK)

2.1.2. Equipment

Avanti J-20 Centrifuge (Beckman Coulter, *High Wycombe, UK*)

Becton Dickinson FACS Calibur (Franklin Lakes, *NJ, USA*)

BioDoc-It UV transilluminator system (Ultra-Violet products- UVP Ltd., *Upland, CA, USA*)

Biofuge Pico microfuge (SORVALL ®, *New Castle, Delaware, USA)* BioPhotometer (Eppendorf, *Hamburg, Germany*)

Centrifuge Universal Legend RT (SORVALL ®)

Dyad thermocycler (MJ Research Inc., *St Bruno (Quebec), Canada*)

Flourescence light Microscope, Eclipse TS100 (Nikon, *Melville, NY, USA*)

Light Microscope, TM (Nikon)

Micro Centrifuge Pico Biofuge (SORVALL ®)

Multitron shaking incubator (Infors AG, *Bottmingen/Basel, Switzerland*)

Nano-Drop 1000D ((Thermo Fisher Scientific, Loughborough, UK)

Shaking Multitron Incubator (Infors AG)

Thermanox® coverslips (Nalge Nunc Int. *Penfield, New York, USA*)

UV transilluminator (UVP Ltd)

2.1.3. General chemicals and reagents

2.1.3.1. General chemicals and reagents used in cloning procedures

Ampicillin (Sigma-Aldrich, *St.Louis, Missouri, USA*)

Bovine serum albumin (Sigma-Aldrich)

DEPC treated water (Life Technologies Ltd, Paisley, UK)

DNA Rehydration Solution (Promega)

Ethanol (BDH Merck Ltd., *Poole Dorset, UK*)

Ethidium bromide (Sigma-Aldrich)

Glycerol (Sigma-Aldrich)

Isopropanol (BDH Merck Ltd.)

$MgCl_2$, 25mM (Promega)

NEB buffer 1 (NEB®)

NEB buffer 2 (NEB®)

NEB buffer 3 (NEB®)

NEB buffer 4 (NEB®)

Phenol (Sigma-Aldrich)

Phenol: chloroform : isoamyl Alcohol 25 : 24 : 1, v/v (Life Technologies Ltd, Paisley, UK)

Restriction endonucleases: (NEB®)

Taq Polymerase (Promega)

Tris Base (Sigma-Aldrich)

TE buffer solution (Life Technologies Ltd, Paisley, UK)

Tris Base (Sigma-Aldrich)

Tris-HCl (BDH Merck Ltd.)

Tween-20 (Sigma-Aldrich)

Deoxynucelotide triphosphates (dNTPs), 100mM: (GE Healthcare Bio-Science Corp., *Chalfont St. Giles, UK*)

2.1.3.2.Reagents used in tissue culture experiments

Dimethyl sulphoxide – DMSO (Sigma-Aldrich)

Fetal calf serum (PAA laboratories Ltd, Somerset, UK)

Dulbecco's Modified Eagle's Medium (DMEM) + Glutamax (PAA)

Formaldahyde –PFA (Sigma-Aldrich) Optimem® (Life Technologies Ltd)

Phosphate buffered saline (PBS) (Sigma-Aldrich)

Phosphate buffered saline + Calcium and Magnesium (PBS +) (Sigma-Aldrich)

Polyethylenemine (PEI) (Sigma-Aldrich)

Polybrene (Sigma-Aldrich)

Tripsin TrypLE express (Life Technologies Ltd)

Trypan blue (Sigma-Aldrich)

2.1.4. Bacterial strains

DH5α Chemically competent *E.coli* (InvitrogenTM)

GeneHogs Electrocompetent *E.coli* (InvitrogenTM)

TOP10 One Shot® Electrocompetent *E.coli* (InvitrogenTM)

XL10 GOLD chemically competent *E.coli* (Stratagene®)

2.1.5. Prepared solutions

All general laboratory reagents were minimally of Analar Grade.

Luria Bertani (LB) medium: 10 g Tryptone, 5 g Yeast extract and 10 g NaCl were dissolved in 1 litre of distilled and deionized water, and then autoclaved on liquid cycle for 20 min at 15 psi. The solution was then allowed to cool to 55°C, and made 50 μg/mL with Ampicillin. The medium was then stored at 4°C.

LB agar plates: 20 g agar per litre of LB broth. LB broth or LB agar media were supplemented with 100 mg ampicillin per litre for the production of transformed bacteria selective media.

TE buffer – 10x (Tris-EDTA buffer):

14 g EDTA (electrophoresis grade) (Sigma-Aldrich) plus 30 g Tris base (Sigma-Aldrich) made up to 1 litre with ddH_20, pH to 8 with HCl.

200 mM methanol (BDH Merck Ltd) in ddH_20.

3M Sodium acetate adjusted to pH5.2 with glacial acetic acid, and sterilised by autoclaving.

50x Tris-Acetate-EDTA (TAE) 50x electrophoresis buffer:

242 g Tris base (electrophoresis grade) (BDH Merck Ltd), 57.1 mL glacial acetic acid (BDH Merck Ltd) and 100 mL 0.5M EDTA (pH 8) made up to 1L with ddH_20.

2.2. Methods

2.2.1. Bacterial cultures

Commercially available competent *E coli* DH5α (Life Technologies Ltd) was used for plasmid transformation in accordance with the manufacturer's instructions.

Standard aseptic techniques were used in bacteria handling. All cultures were carried at 37°C for approximately 16 hours. A shaking incubator set at 180-200rpm was use for liquid cultures. Solid cultures on petri dishes and liquid cultures in universal tubes were kept at 4°C for a short period of time (maximum two weeks). A 20%v/v glycerol stock of overnight cultures of clones colonies containing plasmids were stored at -80°C.

2.2.2. Plasmid production

Standard methods were used to prepare and store bacteria containing plasmids harbouring cloned genes. Bacterial stocks were streaked onto agar plates with 100 μg/mL ampicillin and grown overnight at 37ºC. Single colonies were subsequently picked and used to inoculate liquid cultures of LB medium containing 100 μg/mL ampicillin. Small scale, mini-prep cultures were of 5ml and large-scale maxi-prep cultures were 200-500 mL in volume. Cultures were expanded at 37ºC in a shaking incubator overnight. Plasmid extraction from mini-prep and maxi-prep cultures was performed using the appropriate Qiagen kit and eluted in TE buffer as per the manufacturer's protocols. Concentrations

of DNA were then determined using a NanoDrop 1000 Spectrophotometer and stored at -20ºC.

2.2.3. Gel electrophoresis of DNA fragments

Standard electrophoresis in 0.8-2% agarose gels at 4-200V with 1x TAE buffer were used to resolve and analyse DNA fragments from restriction enzyme digests.

Gels and tank buffer contained 0.005% ethidium bromide for UV visualisation of DNA bands. The size of DNA fragments was estimated by comparison to 100bp and 1kb commercially available size ladders (NEB).

2.2.4. Restriction enzyme digestion analysis of plasmid DNA

Digests were performed in reaction mixtures containing:

1-40 µg DNA, depending on individual experimental requirements,
10% NEB buffer (NEB Buffer 1 – 4, as appropriate for the enzymes),
10% 10x BSA (if required by any enzyme),
1 unit/µg DNA of appropriate restriction endonuclease(s),

made up to the required volume in ddH_2O such that the restriction enzyme addition is not more than 10% of the total.

2.2.5. Lentiviral vector production and quantification

2.2.5.1.Transfection of virus producer cells

Human embryonic kidney 293T (HEK293T) cells were maintained and manipulated in DMEM medium supplemented with 10% fetal calf serum (FCS), 1% 200 mM L-glutamine (Life Technologies Ltd) and 10 μg/mL each of penicillin and streptomycin (Life Technologies Ltd) at 37^0C and 5% CO_2 atmosphere.

T162 flasks (Corning, Flintshire, UK) were seeded with $2x10^7$ HEK293T cells such that the cultures were 80-90% confluent at the time of transfection. The following amounts of DNA were added to 5 mL of Optimem (Life Technologies Ltd) per T162 flask and before being filtered through a 0.2 μm acetate membrane syringe filter (Thermo Scientific UK):

Vector construct: 50 μg

Envelope plasmid (pVSV-G): 17.5 μg

Packaging plasmid (p8.74): 32.5 μg

A 1μL aliquot of a 10mM stock of PEI was also added to 5mL of Optimem and likewise filtered through a 0.2μm filter. DNA and PEI solutions were then mixed in a 1:1 ratio and incubated at room temperature for 20 minutes. The cells were washed once with 10 mL Optimem and the 10 mL of PEI/DNA mixture was then added per T162 flask. The cells were then incubated for 4 hours followed by replacement of the transfection medium with 25 mL of

complete HEK293T DMEM culture media. The media was likewise replaced again 24 hours later.

2.2.5.2.Harvesting of lentiviral vector particles

At 48 hours post-transfection the medium was harvested and replaced with 25 mL fresh DMEM so that virus-containing culture supernatant could be harvested for a second time at 72 hours. The harvested medium was filtered into 50 mL Falcon tubes through a 0.2 μm filter and then centrifuged at 4ºC for 24 hours at 3000x*g*. The supernatant was then removed by decantation and the tubes inverted onto absorbent tissue to drain the remaining fluid. A 50 μL aliquot of Optimem was then added to the viral pellet, which was incubated on ice for 20 minutes, taking care that the fluid level in the tubes was over the pellet's location. The viral pellet was then resuspended in the Optimem by repeated gentle pipetting. Aliquots of 5-30 μL were then made and stored at -80C.

2.2.5.3.Viral vector titration

For each viral harvest, 6 wells of a 24-well plate were seeded with 1-2x10^5 HEK293T cells, allowed to settle for 4 hours and transduced with serial dilutions of concentrated viral stock in order to achieve a range of multiplicity of infections (MOI) between 1.0 and 1x10^{-5} times that obtainable from 1 μL of the virus stock. Specifically, 1μL of viral particle suspension, which had been

frozen and thawed once, was added to 2ml of medium in one well; from the remaining stock, 1/10 serial dilutions were then carried out, such that additional wells were administered with $1x10^{-1}$, $1x10^{-2}$, $1x10^{-3}$, and $1x10^{-4}$ µL of stock. To the 1 mL of medium in each well, a further 1mL of media was added containing 16 µg/mL of polybrene to give a final concentration of 8µg/mL. The sixth well was left untransduced to include as a negative control. Calculation of the MOI of 1 mL of stock was then possible by analysis via flow cytometry to discover what percentage of the HEK293T cells in a given well were transduced by the dilution applied and scaling up as appropriate.

2.2.6. Cell harvesting for flow cytometry analysis

Adherent cells were detached by the application of Tryple Red reagent (Life Technologies Ltd) after washing cells with PBS and subsequent neutralisation of the Tryple Red with serum-containing medium. Cells were then transferred to 14 mL or 5 mL tubes and centrifuged for 5 minutes at 1000rpm. The supernatant was then removed and the cell pellet resuspended in desired medium and volume for further manipulation. In the case of analysis by flow cytometry, the pellet was resuspended by adding 1 mL of 4% formaldehyde in PBS to fix cells and shaking the tube vigorously; samples were then stored at 4^0C in the dark until analysis.

2.2.7. Quantification of viral titre by flow cytometry

Samples of at least $2x10^5$ fixed cells were analysed for GFP fluorescence by flow cytometry. Only live cells within the total cell population were considered via analysis of size (measured by forward scatter) versus granularity (measured by side scatter) and then detecting light emission at 525 nm (in the colour spectrum range FL1) versus light emission at 575 nm (in the colour spectrum range FL2) with the aid of control untransduced cells to gate for positive GFP signal. This provides a measure of the percentage of cells in the sample that were both alive and expressing GFP. The percentage of GFP positive cells in each sample was related to the concentration of virus that the sample had received and scaled up to give the number of cells transduced by that amount of viral stock, thereby permitting the calculation of MOI.

Values of between 1% and 20% were preferred for this calculation, as samples with over 20% GFP positive cells were likely to harbour more than one integration event per cell and thus cause the underestimation of viral titre, whereas those showing less than 1% expressing GFP at detectable levels were not useful due to being too likely to be distorted by background noise.

2.2.8. Maintenance and differentiation of P19 cells

Murine embryonic carcinoma P19 cells were maintained and manipulated in Dulbecco's modified Eagle's medium (DMEM; GE Healthcare Life Sciences) supplemented with 2 mM L-glutamine (Life Technologies Ltd), 1% non-

essential amino acids (Life Technologies Ltd), 10% fetal calf serum and 10 μg/ml each of penicillin and streptomycin in a humidified 5% CO_2 atmosphere. Differentiation of P19 cells down the ectodermal lineage into neuronal cells was performed by inducing cell aggregate formation to form embryoid bodies on non-adhesive bacterial-grade Petri dishes in the presence of 1 μM of ***all-trans***-retinoic acid (RA; SigmaAldrich Ltd) at a density of 1×10^5/mL in DMEM containing 5% FCS. After 2 days of RA treatment, embryoid bodies were dissociated with 0.25% trypsin, 0.53 mM EDTA (Life Technologies Ltd), and re-plated in poly-D-lysine-coated (0.1 mg/mL) tissue culture dishes and allowed to differentiate for a further 2-8 days in the presence of 10 μM cytosine arabinoside to enrich for non-dividing neuronal cells.

2.2.9. Maintenance and differentiation of F9 cells

F9 embryonal carcinoma cells were cultured in DMEM supplemented with 10% FCS, 100 units/mL penicillin, 100 μg/mL streptomycin, and 2mM L-glutamine, in a humidified 5% CO_2 atmosphere on gelatin (0.1%) coated flasks. Differentiation of F9 cells into parietal endoderm cells was performed by inducing cell embryoid body formation in non-adhesive bacterial-grade Petri dishes in the presence of 50 nM RA at a density of 1×10^5/mL in DMEM/F12 containing 5% FCS. After 5-6 days of RA treatment, embryoid bodies were dissociated with 0.25% trypsin, 0.53mM EDTA and re-plated in poly-D-lysine-coated (0.1 mg/mL) tissue culture dishes and allowed to differentiate for a further 2-6 days with 10 μM cytosine arabinoside in the culture medium to enrich for non-dividing endodermal cells.

2.2.10. Maintenance and differentiation of murine ESCs

Murine CCE embryonic stem cells (ESCs) were maintained in an undifferentiated state in DMEM medium (Sigma-Aldrich) supplemented with leukemia inhibitory factor (LIF, 103U/mL; Chemicon, Millipore (U.K.) Limited, Watford, Herts, UK), 1% L-glutamine, 20% batch tested FBS (Summit), 1% penicillin/streptomycin (Sigma-Aldrich) and $1.5x10^{4}$ M monothioglycerol (MTG, Sigma-Aldrich). Cells were cultured in 6 well plates coated with 0.1% gelatin in PBS (Sigma-Aldrich). Once 80% confluent, cell cultures were trypsynised with 0.25% Trypsin-EDTA (Sigma-Aldrich) and split 1:2 or 1:3.

Prior to differentiation to embryoid bodies (EBs), cells were passaged and cultured in IMDM medium (Life Technologies Ltd) supplemented with the same substances as with the DMEM media.

In order to initiate EB formation, cells were trypsinised and re-seeded into non-adherent, uncoated 60 mm Falcon dishes (Thermo Fisher Scientific) at $0.375x10^{6}$ cells/dish in serum-free media (SFD) in the absence of LIF but instead containing the following supplements: N2 (Life Technologies), F12 (Life Technologies), B27 (Life Technologies), IMDM, 1% L-glutamine, 1% penicillin/streptomycin, and 10% BSA/PBS (Sigma-Aldrich). Prior to adding media to the cells, SFD was further supplemented with 500 mg/mL L-ascorbic acid, MTG (13 μl/mL) and 1% L-glutamine.

2.2.11. Vector copy number estimation analysis

To confirm whether reduced eGFP expression over time was due to either loss of vector-positive cells from the culture or LV silencing, quantitative PCR (qPCR) was performed on total genomic DNA to determine average vector copy number (VCN) per cell at each time point of the period of culture after transduction. A cell clone harbouring a single A2UCOE-eGFP LV copy was isolated by limiting dilution of bulk transduced P19 and F9 cells and used as a reference in the VCN qPCR determination.

2.2.12. Immunofluorescence counting of transduced differentiated P19 and F9 cells

After P19 and F9 cell differentiation to embryoid bodies, the cells were removed from the bacterial culture dishes and transferred into 6-well plates, within which were placed laminin coated glass coverslips. Upon completion of the period of differentiation, immunofluorescence staining was conducting as follows.

Cells on the laminin coated glass coverslips were washed three times with PBS and then fixed with 4% paraformaldehyde in PBS twice for 10 min. The fixed cells were then treated with blocking buffer (1 x PBS/0.1% Triton X-100) for 30

min. The slides were washed with 5% normal serum in PBS three times. Primary antibodies anti-β-tubulin III antibody (Tuj1) for P19 cells and anti-mouse Oct3/4 for F9 cells, were prepared in PBS + 1% gelatin at a 1/200 and 1/500 dilution, respectively. The slides with fixed cells were then incubated with appropriate primary antibodies for 1hr in a humidified chamber at room temperature in the dark. Cells were then washed three times in PBS, 5 min each wash. The secondary antibody AlexaFluor 488 goat anti-mouse was then applied for 1 hour at room temperature in a humidified chamber. The cells were then again washed three times with PBS. Finally, slides were incubated with 0.1-1 μg/mL Hoechst and DAPI for 1 min and the coverslips sealed on microscope slides with nail polish to prevent drying and movement under microscope. The samples were stored in dark at +4°C.

Immunofluorescence counting of the cells was conducted under confocal microscopy. Neuronal differentiated P19 cells were detected by secondary antibody red fluorescence of β-tubulin III antigen and endodermal differentiated F9 cells by Oct3/4 detection. Cell nuclei were highlighted by the blue DAPI stain.

Chapter 3
Results

Evaluating novel candidate UCOEs with a dual, divergently transcribed housekeeping gene promoter configuration

Chapter 3

Evaluating novel candidate UCOEs with a dual, divergently transcribed housekeeping gene promoter configuration

3.1 Introduction and background

The proposed mechanism, by which the CBX3-HNRPA2B1 UCOE (A2UCOE) brings about a dominant chromatin opening function, is a two-component system. First, a large (2.7kb) CpG island in association with an extended methylation-free region (5kb) with an inherent open chromatin structure and second, chromatin remodeling via dual divergent transcription from the innate *CBX3* and *HNRPA2B1* promoters (Allen and Antoniou 2007). The work presented in this chapter aimed to further test this two-component mechanism of A2UCOE function by evaluating other dual divergent transcription systems based on housekeeping genes as in the case of *CBX3* and *HNRPA2B1*.

In addition, it was previously reported that the placement of the core A2UCOE upstream of heterologous promoters would confer stability of expression in an orientation specific manner; that is, when the CBX3 end of the A2UCOE was placed next to the heterologous ubiquitous SFFV (Zhang, Frost et al. 2010) or tissue-specific MRP8 (Brendel, Müller-Kuller et al. 2012) promoter elements. It was suggested that the reason for this orientation-bias in A2UCOE function when linked to heterologous promoters in certain contexts may be due to the relatively weak transcriptional activity of the *CBX3* promoter compared to that of *HNRPA2B1*. Thus, divergent transcription from the stronger *HNRPA2B1*

promoter would provide a better barrier against epigenetic (DNA methylation, histone modification) mediated silencing. We, therefore, sought to construct UCOEs with a dual divergent transcriptional configuration but which may function equally in either orientation when linked to heterologous promoters.

The lentiviral vector (LV) constructs tested here were conceived and built by a previous member of the Antoniou lab (Thalia Vlachou). A search of the literature (She, Rohl et al. 2009) and bioinformatics (UCSC Genome Browser) analysis of housekeeping genes was undertaken to identify gene pairs based on the criteria of high, uniform expression across a variety of normal human tissues. Based on this search the *SETD3-CCNK* housekeeping gene pair was chosen as a good novel UCOE candidate. *SETD3* and *CCNK* constitute a dual divergently transcribed pair of genes and which are both relatively uniformly expressed across 40 different human tissues and cell types. In order to further test the hypothesis that the orientation bias of the A2UCOE when linked to heterologous promoters was due to the weak transcriptional activity of *CBX3*, an artificial UCOE was constructed whereby the HNRPA2B1 component of the A2UCOE was linked to the single promoter-CpG island of the *RPS11* housekeeping gene in a divergently transcribing configuration. The *RPS11* element was again chosen as this gene is expressed at a uniformly high level across different human tissues (She, Rohl et al. 2009).

3.2 LV used in this study

Note: the LV constructs SEW (SFFV-eGFP-WPRE), 1.5A2UCOE-SEW, SET-CCN-SEW, CCN-SET-SEW, B1-RPS-SEW, RPS-B1-SEW described below in Figure 1 and used in this study were generated by previous members of the Antoniou lab (ThaliaVlachou, Olivia Bales).

3.2.1. SEW (SFFV-eGFP-WPRE)

SEW was derived from pHR'SIN-cPPT-SEW (Demaison, Parsley et al. 2002) with the addition of a restriction enzyme multi-cloning polylinker inserted at the unique *Eco*RI site 5' of the SFFV promoter (Figure 3.1). pHR'SIN-cPPT-SEW is a second generation HIV-1 based SIN transfer vector, in which the reporter gene eGFP is located downstream of the SFFV promoter (Figure 3.1). A WPRE element is also present downstream of eGFP. The plasmid vector backbone contains a *bla* gene, which is an AmpR marker.

SEW (SFFV-eGFP-WPRE)

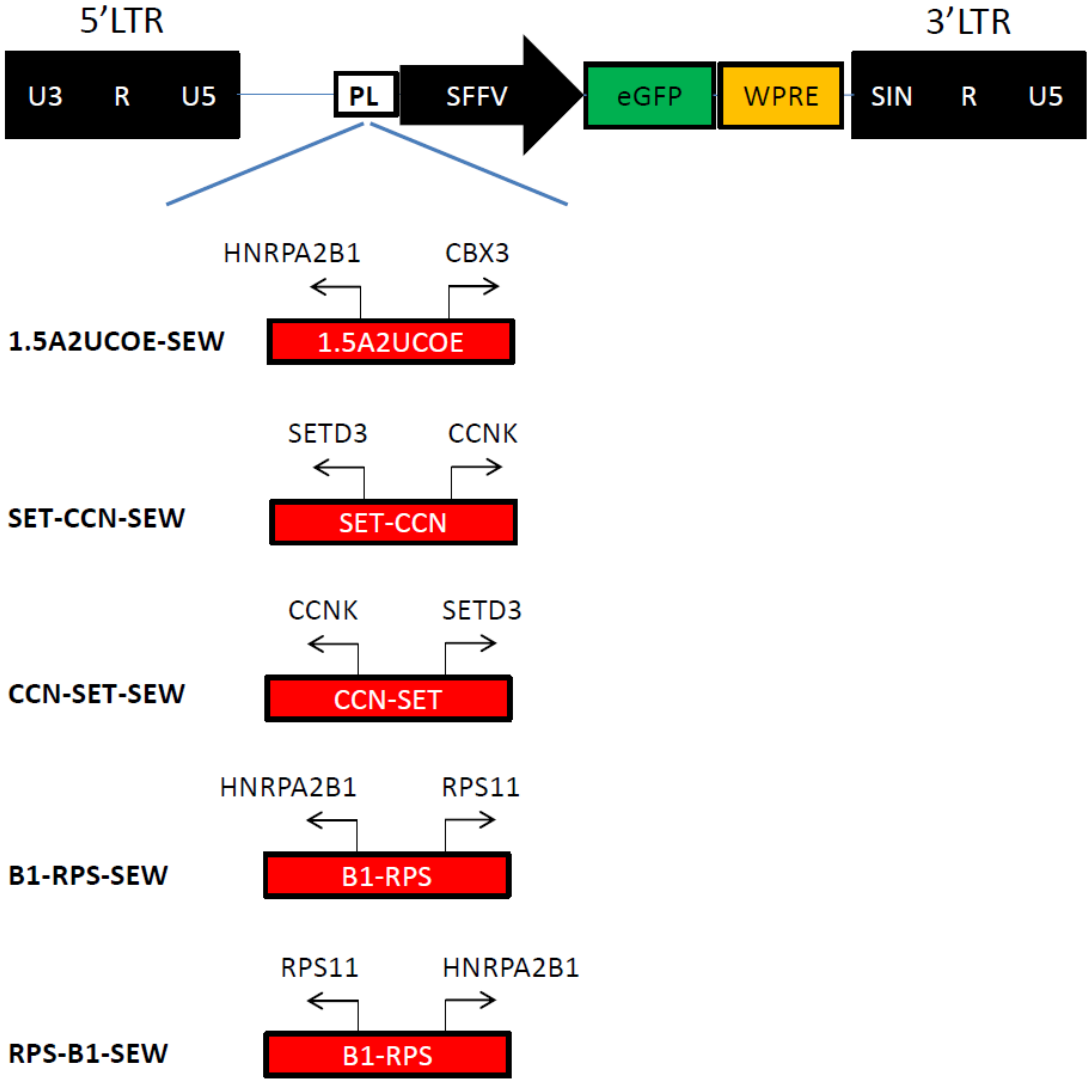


Figure 3.1. Illustration of the novel candidate UCOE and control lentiviral vectors. A standard self-inactivating (SIN) lentiviral vector backbone containing a spleen focus forming virus (SFFV) promoter driving expression of an enhanced green fluorescent protein (eGFP) reporter gene with downstream woodchuck hepatitis virus post-transcriptional regulatory element (WPRE) designated as SEW acted as the starting point for the construction of test UCOE vectors. The positive control 1.5kb HNRPA2B1-CBX3 UCOE (1.5A2UCOE) was inserted into a polylinker (PL) cloning site upstream of the SFFV promoter to give 1.5A2UCOE-SEW. Novel candidate UCOEs consisting of dual divergently transcribed promoter regions from the human SETD3-CCNK locus and artificially constructed RPS11-HNRPA2B1 were inserted into the PL in both orientations to give SET-CCN-SEW/CCN-SET-SEW and B1-RPS-SEW/RPS-B1-SEW constructs, respectively. LTR: long terminal repeat.

3.2.2. A2UCOE-SEW

1.5A2UCOE-SEW was generated by subcloning of the 1.5kb core A2UCOE fragment (Williams, Mustoe et al. 2005) within the polylinker at the unique *Bst*X1-*Eco*RI sites upstream of the SFFV promoter of SEW (Figure 1). The orientation of the 1.5kb A2UCOE in this vector places the *CBX3* end of the element next to SFFV and thus *HNRPA2B1* transcribing in a divergent direction (Figure 3.1). This is the orientation found to confer stability of expression from the SFFV promoter following transduction into P19 cells (Zhang, Frost et al. 2010).

3.2.3. SET-CCN-SEW and CCN-SET-SEW

SET-CCN-SEW and CCN-SET-SEW were derived by subcloning of the SETD3-CCNK CpG island / dual divergently transcribed promoter region in the *Pac*I site of the polylinker upstream of the SFFV promoter of SEW in both orientations (Figure 3.1).

3.2.4. B1-RPS-SEW and RPS-B1-SEW

The single promoter CpG island of *RPS11* was linked in a divergent transcribing orientation to the *HNRPA2B1* promoter derived from the A2UCOE with CBX3 deleted. B1-RPS-SEW, RPS-B1-SEW are derived by inserting the artificially combined dual promoter, CpG island HNRNPA2B1-RPS11 element

into the *Pac*I site of the polylinker of SEW upstream of the SFFV promoter (Figure 3.1).

3.3 Restriction enzyme digestion confirmation of LV plasmid preparations

The identity and integrity of large-scale maxi-preparations of the plasmid LV to be tested (Figure 3.1) was via restriction enzyme digestion and resolution by agarose gel electrophoresis. All constructs gave the expected banding pattern confirming their correct status and quality (Figure 3.2).

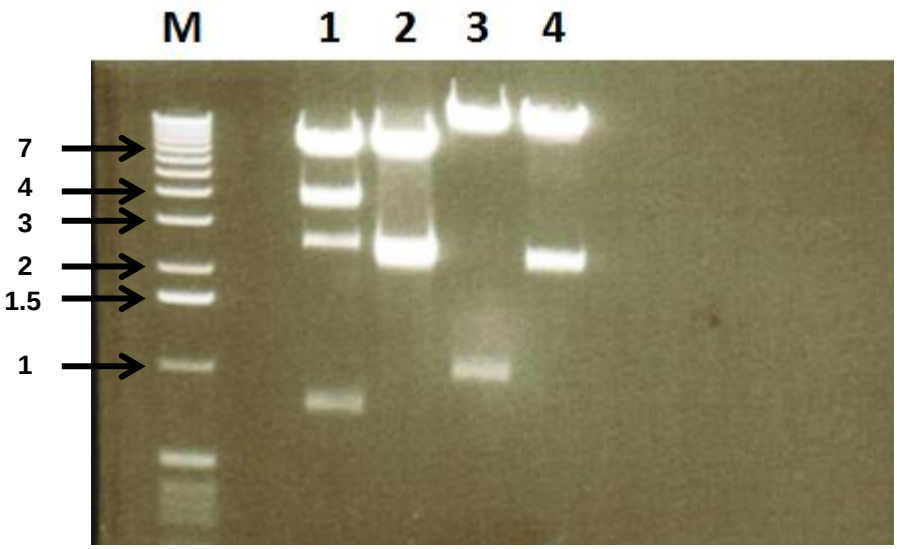


Figure 3.2. Confirmation of identity and integrity of novel candidate UCOE lentiviral vector plasmid preparations. Maxi-preparations of novel candidate UCOE (Figure 3.1) plasmid preparations were digested with appropriate restriction enzymes and resolved by agarose gel electrophoresis. Lane 1: SET-CCN-SEW digestion with *Bsp*EI/*Bam*HI to give 795bp, 2725bp and 4520bp diagnostic bands. Lane 2: CCN-SET-SEW digestion with *Bsp*EI/*Bam*HI to give a 2358bp diagnostic band. Lane 3: B1-RPS-SEW digestion with *Bsi*WI/*Bam*HI to give a 963bp diagnostic band. Lane 4: RPS-B1-SEW digestion with *Bsi*WI/*Bam*HI to give a 2158bp diagnostic band. M: molecular size markers in kb.

3.4 Lentiviral vector titration in HEK293T cells

Cells were collected for analysis by flow cytometry three days after the transduction of $2x10^5$ HEK293T cells with serial dilutions of harvests at both 2 and 3 days post-transfection from each production run for all LVs: SEW (SFFV-eGFP-WPRE), 1.5A2UCOE-SEW, SET-CCN-SEW, CCN-SET-SEW, B1-RPS-SEW, RPS-B1-SEW (Table 3.1).

LV	Volume (µL) of viral vector stock			
	2 µL	0.2 µL	0.02 µL	0.002 µL
	Percentage eGFP+ cells 1st / 2nd viral harvest			
SEW	62.21 / 74.32	11.38 / 13.57	1.14 / 1.46	0.41 / 0.52
1.5A2UCOE-SEW	66.23 / 54.6	8.5 / 7.12	0.82 / 1.01	0.21 / 0.3
SET-CCN-SEW	78.24 / 64.26	14.86 / 11.36	1.84 / 1.23	0.00 / 0.00
CCN-SET-SEW	24.68 / 16.78	1.65 / 1.47	0.00 / 0.00	0.00 / 0.00
B1-RPS-SEW	36.81 / 42.11	8.05 / 9.62	1.21 / 1.32	0.23 / 0.33
RPS-B1-SEW	8.66 / 4.65	0.64 / 0.55	0.00 / 0.00	0.00 / 0.00

Table 3.1: Titration of lentiviral vector (LV) preparations. A $2x10^5$ aliquot of HEK293T cells were transduced with 0.002-2 µL of a given LV stock from the first and second harvest of virus from culture supernatant during the time of preparation. Analysis of eGFP-positive cells at 3-days post-transduction was by flow cytometry. LV constructs are as described in Figure 3.1.

3.5 Calculation of LV titre

Viral titre (Table 3.2) was calculated based on the dilution of each harvest that gave a percentage of eGFP-positive cells between 1% and 10%. Pools of cells with percentages of eGFP-positive cells higher than 10% are very likely to contain multiple integrations of the vector, while very low scores could be false positives.

On day one, $2x10^5$ 293T cells were seeded in each 6-well. Thus, the percentage of eGFP positive cells from flow cytometry analysis reflects the percentage of the initial cell population that was successfully transduced and hence the number of infectious units added in the specified well. The dilution factor or the volume of the lentivirus preparation that was used is known, so the calculation of the number of infectious units per mL from that point onwards is straightforward.

LV	Titre (iu/mL)	
	1st harvest	2nd harvest
SEW	1.14×10^8	1.46×10^8
1.5A2UCOE-SEW	8.5×10^7	7.12×10^7
SET-CCN-SEW	1.11×10^8	1.52×10^8
CCN-SET-SEW	1.65×10^7	1.47×10^7
B1-RPS-SEW	8.05×10^7	9.62×10^7
RPS-B1-SEW	8.66×10^6	4.65×10^6

Table 3.2: Lentiviral vector (LV) titres. Viral titre as infectious units per ml (iu/mL) was calculated based on the dilution of 1st and 2nd harvest of virus that gave a percentage of eGFP-positive HEK293T cells between 1-10% (Table 1). LV constructs are as described in Figure 3.1.

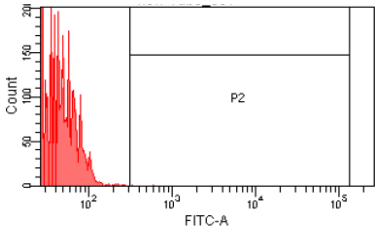
3.6 Functional assay of candidate UCOEs in undifferentiated P19 and F9 cells

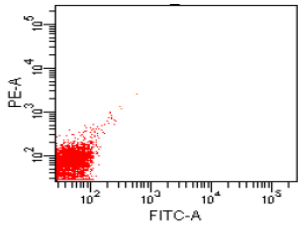
P19 and F9 cells were transduced with the generated LVs at a MOI of 5 or 10, with the intention to start the experiment within the range of 40-60% eGFP-positive cells in all pools. The transduced cells and negative control were then propagated in culture and assayed as a time course by flow cytometry, to determine the percentage of eGFP-positive cells and the mean fluorescence intensity (MFI) for each repeat transduction for each vector. Cell cultures were analysed every 6 days from day 3 post-transduction (Figure 3.3) and extending to 52 days (Figure 3.4). Cell samples were taken at regular intervals for the flow cytometry analysis, to monitor the expression levels of the eGFP reporter gene. In addition, DNA was extracted from cells at each time point for RT-qPCR in order to determine the average vector copy number per cell.

Figure 3.4-A shows the flow cytometry time course results depicting percentage of eGFP+ cells. Although similar initial transduction efficiency was obtained with all vectors (45–60% eGFP+ cells), expression from the SEW (SFFV-eGFP-WPRE), rapidly declined from 58 to 2% positive cells within 17 days. In contrast, the proportion of eGFP+ cells from the 1.5A2UCOE-SEW vector remained completely stable over the 52-day period of culture. Therefore, whilst differences in the percentage eGFP-positive cells transduced with 1.5A2UCOE-SEW was significantly higher when compared with the results obtained for all other vectors, the decrease of eGFP expression of the vectors SET-CCN-SEW, CCN-SET-SEW, B1-RPS-SEW, RPS-B1-SEW is significantly more gradual

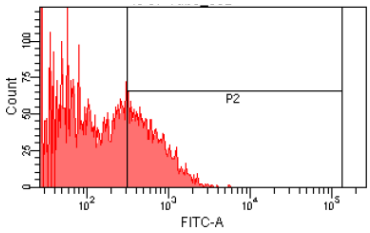
and slower when compared with SEW. The expression of eGFP in SEW transduced cells had dropped almost 80% after 2 weeks post-transduction. Contrastingly, with the exception of 1.5A2UCOE-SEW, the reduction observed in eGFP-positive cells was only 5% to 12% for all other vectors carrying test UCOEs over the same time period. The values of mean fluorescent intensity (MFI) for all vectors paralleled the eGFP expression results (Figure 3.4-B). MFI was stable in the 1.5A2UCOE-SEW transduced cells, unstable in the case of SEW and partially stable in the novel test UCOE constructs.

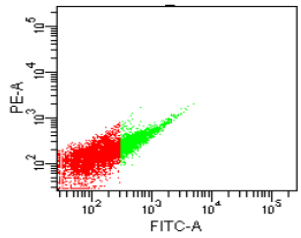
A: NEG



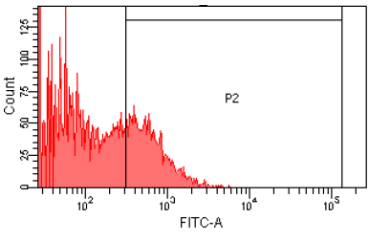


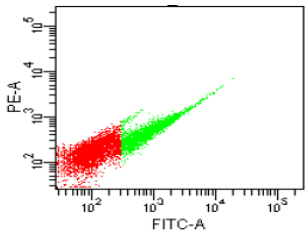
B: SEW



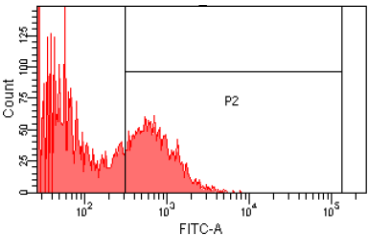


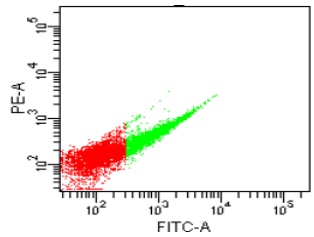
C: 1.5A2UCOE -SEW



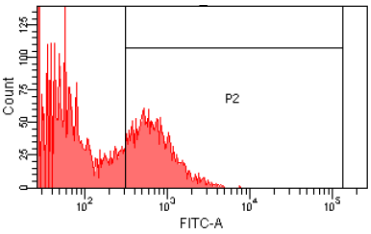


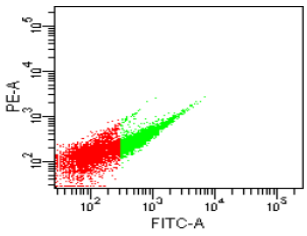
D: SET-CCN -SEW



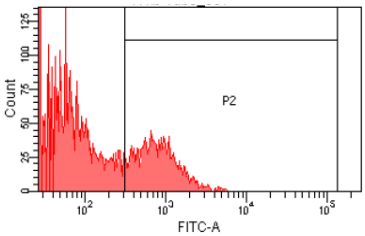


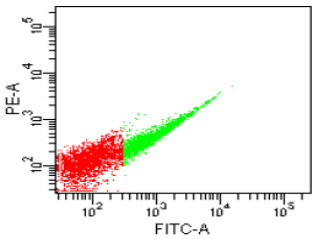
E: CCN-SET -SEW



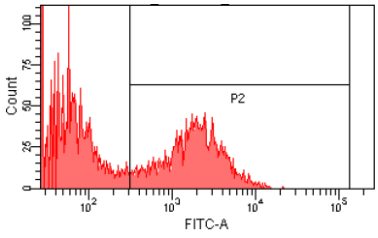


F: B1-RPS -SEW





G: RPS-B1 -SEW



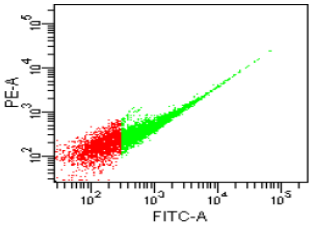
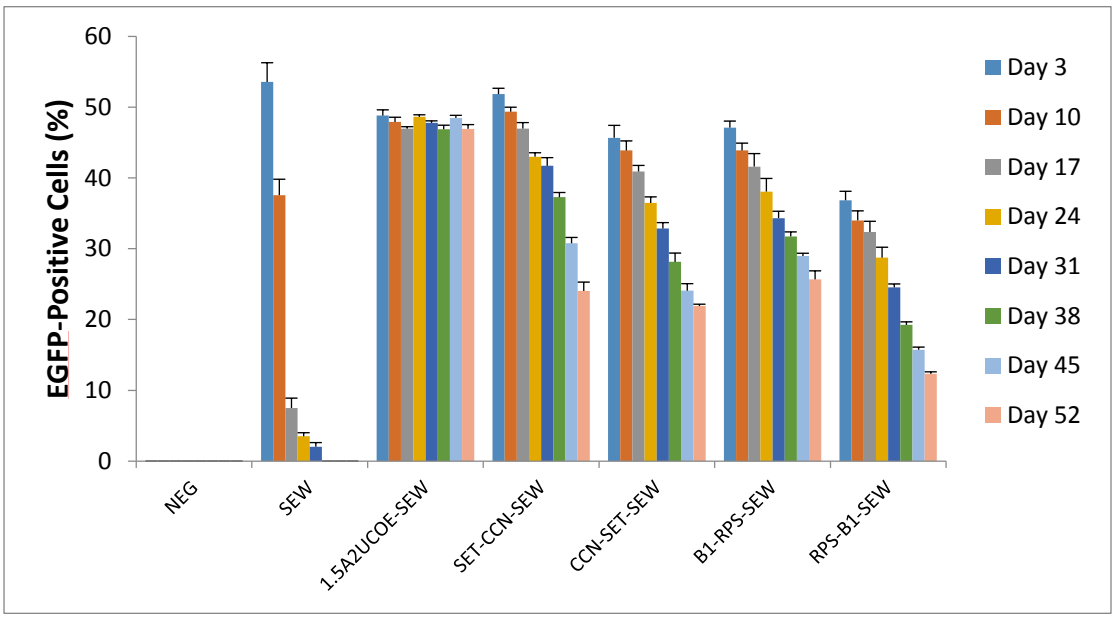


Figure 3.3: Flow cytometry plots of untransduced and transduced P19 cells.

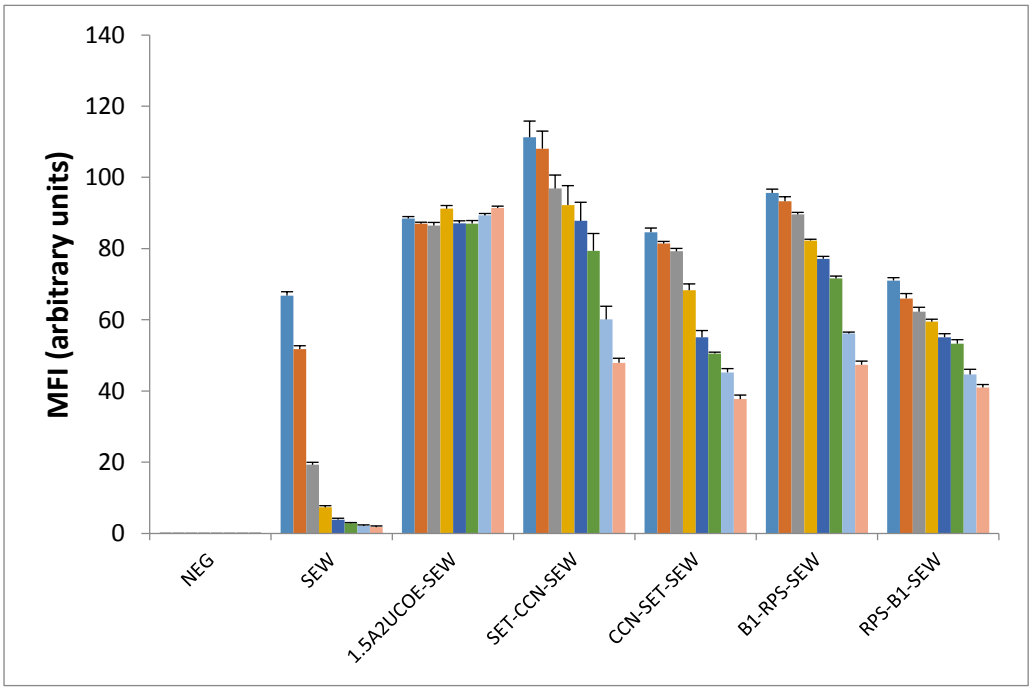
P19 cells were transduced with the test LV (Figure 3.1) at a MOI-3 and cells analysed by flow cytometry 3 days later. GFP-negative cells are shown with red; GFP-positive cells are shown in green. (A) Negative control-untransduced cells. (B) SEW transduced cells. (C) 1.5A2UCOE-SEW transduced cells. (D) SET-CCN-SEW transduced cells (E) CCN-SET-SEW transduced cells (F) B1-RPS-SEW transduced cells (G) RPS-B1-SEW transduced cells.

We next ascertained whether reduced eGFP expression over time was due to the loss of vector-positive cells from the culture. Quantitative RT-Q-PCR was performed on genomic DNA to determine average vector copy number (VCN) at each time point of the period of culture after transduction. Our data show that average VCN per cell remained at a similar level for all vectors at all time points (Figure 3.4-C). This shows that loss of eGFP expression in cells transduced with the SEW and test novel UCOE constructs (Figure 3.4-A) was not due to the loss of vector copies but from SFFV promoter silencing.

A)


EGFP-Positive Cells (%)
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
SET-CCN-SEW
CCN-SET-SEW
B1-RPS-SEW
RPS-B1-SEW
Day 3
Day 10
Day 17
Day 24
Day 31
Day 38
Day 45
Day 52

B)


MFI (arbitrary units)
140
120
100
80
60
40
20
0
NEG
SEW
1.5A2UCOE-SEW
SET-CCN-SEW
CCN-SET-SEW
B1-RPS-SEW
RPS-B1-SEW

C)

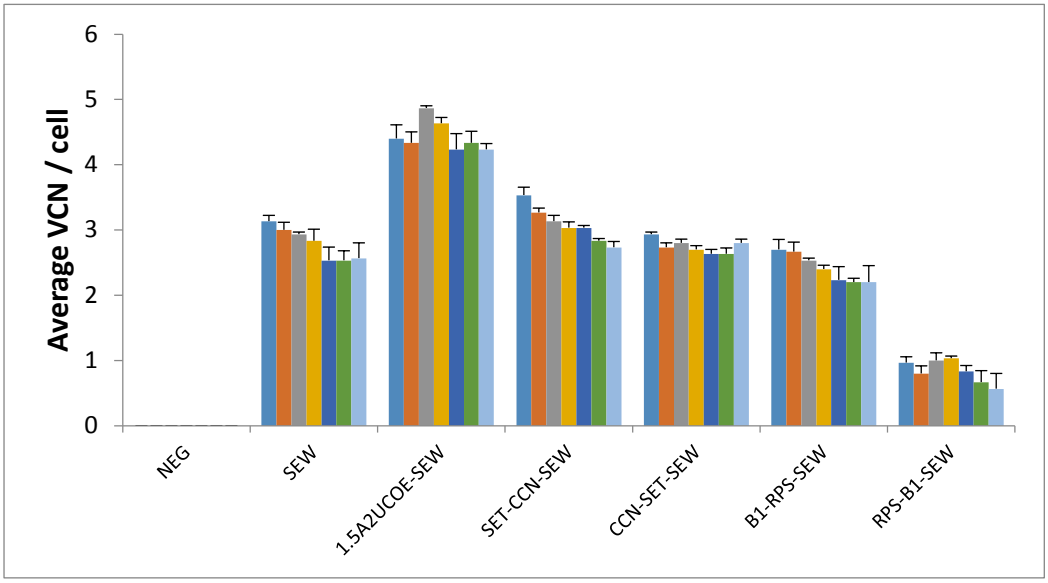
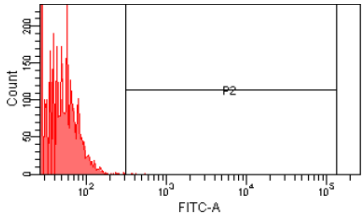
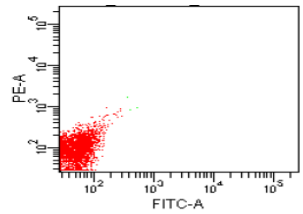


Figure 3.4: Novel candidate UCOEs offer only partial protection against silencing in undifferentiated P19 cells. P19 cells were transduced with novel candidate UCOE SET-CCN-SEW, CCN-SET-SEW, B1-RPS-SEW, RPS-B1-SEW and control SEW and 1.5A2UCOE-SEW lentiviral vectors (Figure 3.1). Cells were analysed by flow cytometry to detect percentage of eGFP reporter gene expressing (eGFP+) cells, mean fluorescence intensity (MFI) and by RT-Q-PCR for average vector copy number (VCN) per cell. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 52 days post-transduction. (**A**) timecourse of percentage eGFP positive cells; (Mean + SEM, n=4; **p<0.01). (**B**) As in (A) but showing MFI; (Mean + SEM, n=4; **p<0.01). (**C**) As in (A)/(B) but average VCN/cell; (Mean + SEM, n=4; **p<0.01).

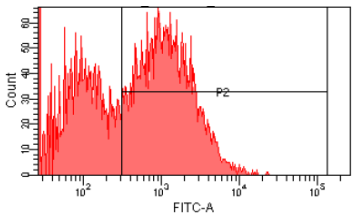
Figure 3.6-A shows the flow cytometry time course results depicting percentage of eGFP+ cells in undifferentiated F9 cells. The results show a parallel activity as with the P19 cells above. Although similar initial transduction efficiency was obtained with all vectors (39–50% eGFP+ cells), expression from the SEW (SFFV-eGFP-WPRE) rapidly declined from 51% to 3% positive cells within 17 days. In contrast, the proportion of eGFP+ cells from the 1.5A2UCOE-SEW vector remained completely stable over the 52-day period of culture. Therefore, whilst differences in the percentage eGFP positive cells transduced with 1.5A2UCOE-SEW was significantly higher when compared with the results obtained for all other vectors, the decrease of eGFP expression of the vectors SET-CCN-SEW, CCN-SET-SEW, B1-RPS-SEW, RPS-B1-SEW is significantly more gradual and slower when compared with SEW. The expression of eGFP in SEW transduced cells had dropped almost 80% after 2 weeks post-transduction. Contrastingly, with the exception of 1.5A2UCOE-SEW, the reduction observed in eGFP positive cells was only 5% to 10% for all other vectors carrying test UCOEs over the same time period. The values of mean fluorescent intensity (MFI) for all vectors paralleled the eGFP expression results (Figure 3.4-B). MFI was stable in the 1.5A2UCOE-SEW transduced cells, unstable in the case of SEW and partially stable in the novel test UCOE constructs.

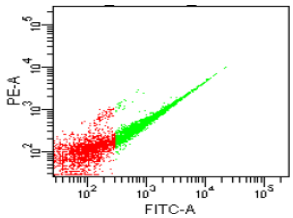
A: NEG



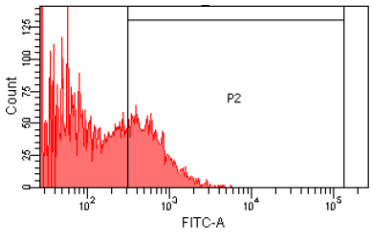


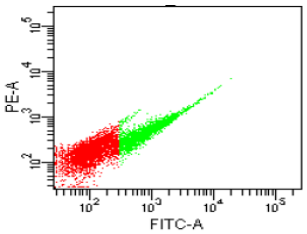
B: SEW



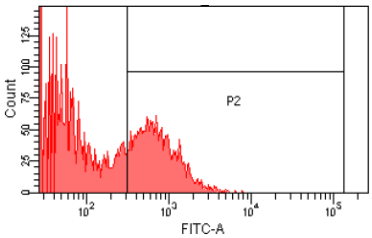


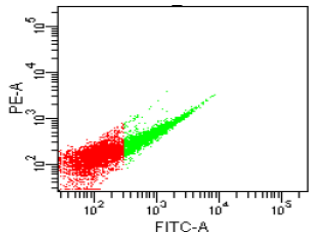
C: 1.5A2UCOE -SEW



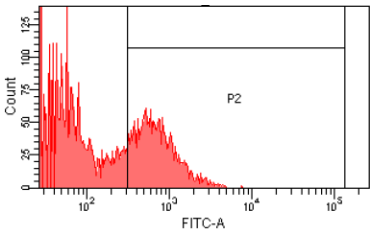


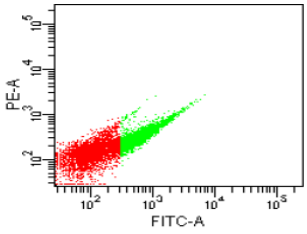
D: SET-CCN -SEW



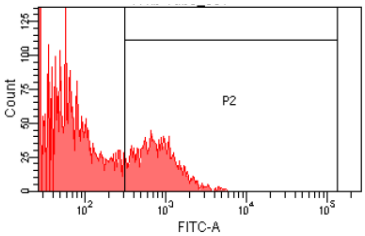


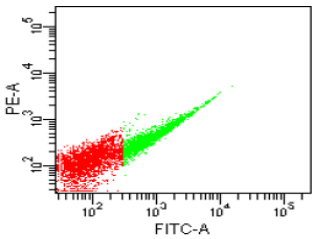
E: CCN-SET -SEW



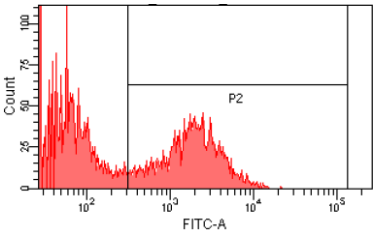


F: B1-RPS -SEW





G: RPS-B1 -SEW



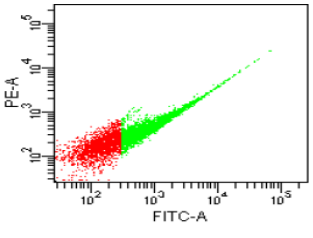
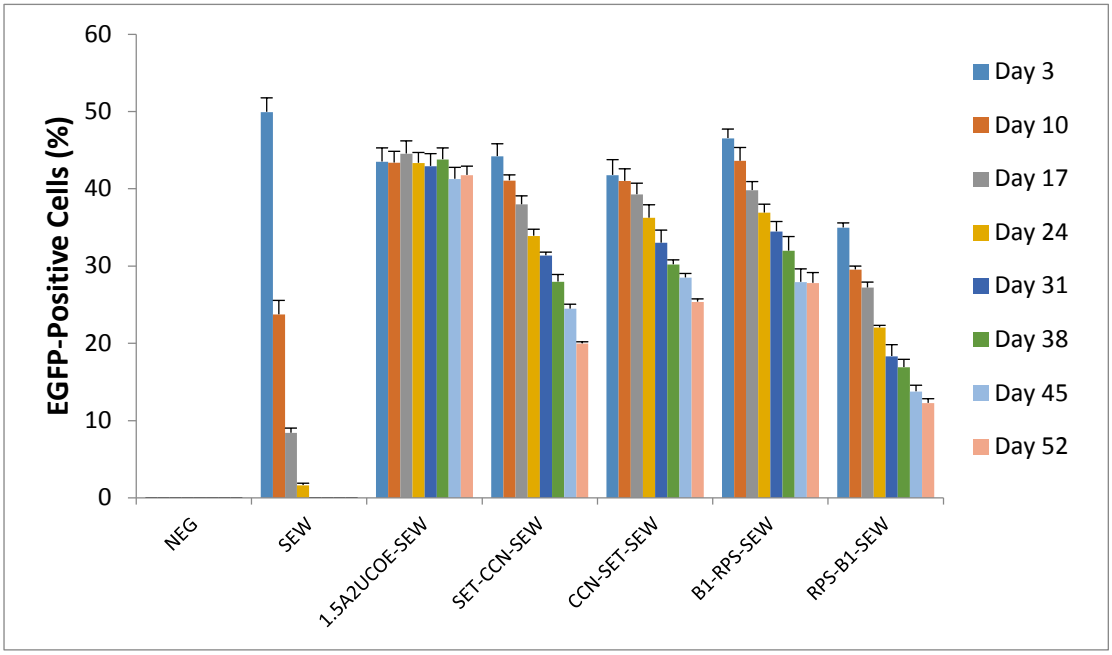


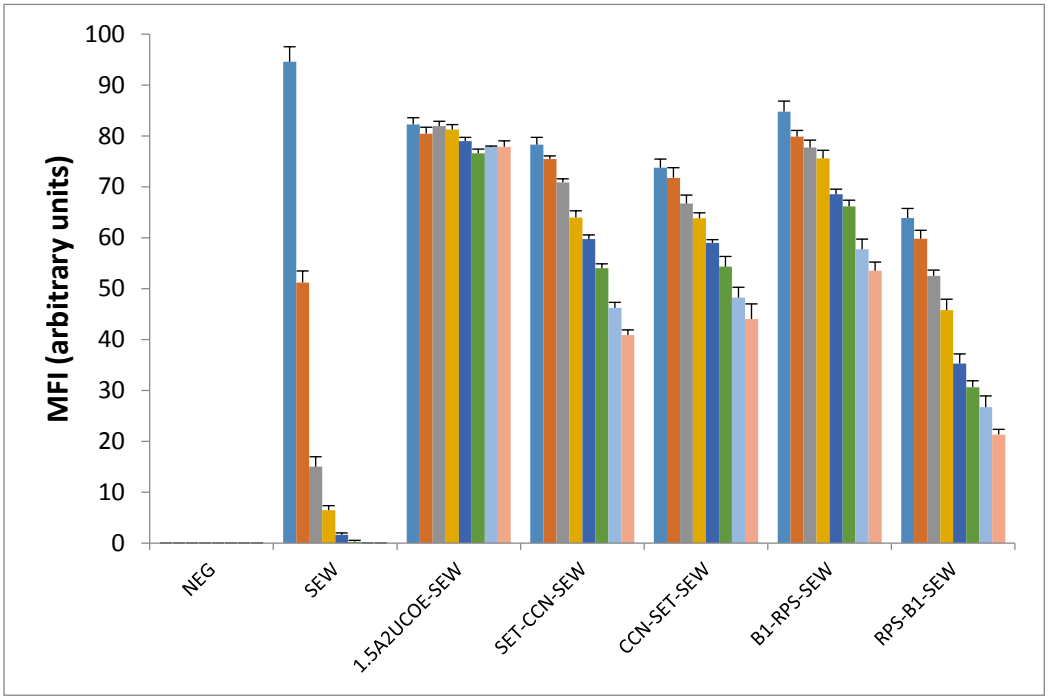
Figure 3.5: Flow cytometry plots of untransduced and transduced F9 cells.

F9 cells were transduced with the control and novel candidate LV (Figure 3.1) at MOI 3 and cells analysed by flow cytometry 3 days later. GFP-negative cells are shown with red; GFP-positive cells are shown in green. (**A**) Negative control-untransduced cells. (**B**) SEW transduced cells. (**C**) 1.5A2UCOE-SEW transduced cells. (**D**) SET-CCN-SEW transduced cells. (**E**) CCN-SET-SEW transduced cells. (**F**) B1-RPS-SEW transduced cells. (**G**) RPS-B1-SEW transduced cells.

A)


EGFP-Positive Cells (%)
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
SET-CCN-SEW
CCN-SET-SEW
B1-RPS-SEW
RPS-B1-SEW
Day 3
Day 10
Day 17
Day 24
Day 31
Day 38
Day 45
Day 52

B)


MFI (arbitrary units)
100
90
80
70
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
SET-CCN-SEW
CCN-SET-SEW
B1-RPS-SEW
RPS-B1-SEW

C)

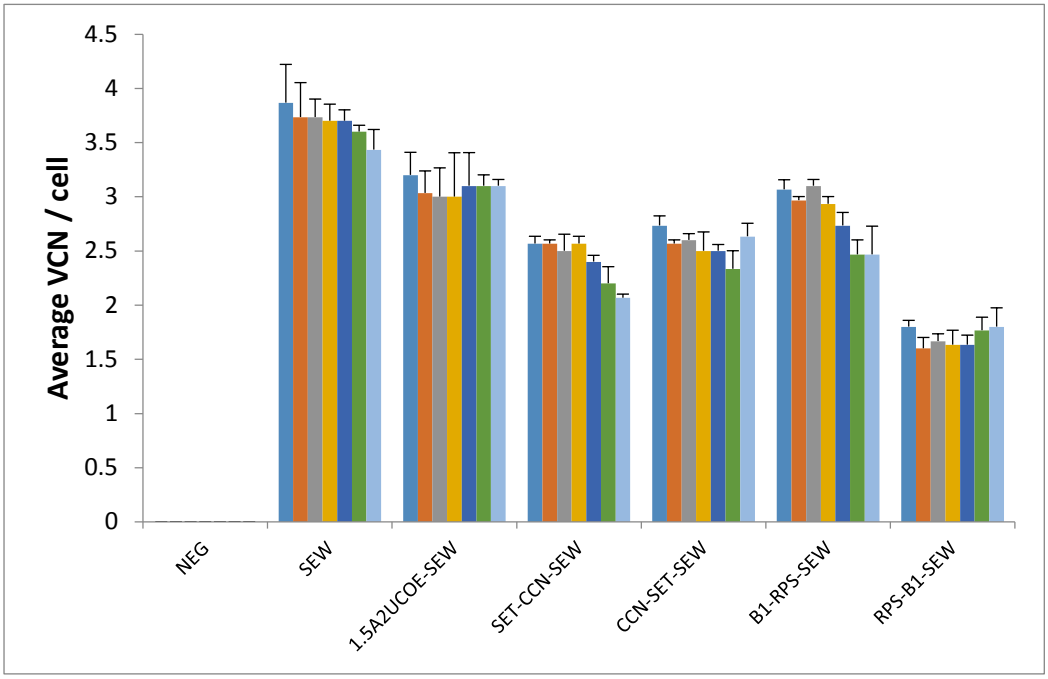


Figure 3.6: Novel candidate UCOEs offer only partial protection against silencing in undifferentiated F9 cells. F9 cells were transduced with novel candidate UCOE SET-CCN-SEW, CCN-SET-SEW, B1-RPS-SEW, RPS-B1-SEW and control SEW and 1.5A2UCOE-SEW LVs (Figure 3.1). Cells were analysed by flow cytometry to detect percentage of eGFP reporter gene expressing (eGFP+) cells, mean fluorescence intensity (MFI) and by real-time Q-PCR for average vector copy number (VCN) per cell. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 52 days post-transduction. (**A**) timecourse of percentage eGFP positive cells; (Mean + SEM, n=4; **p<0.01). (**B**) As in (A) but showing MFI; (Mean + SEM, n=4; **p<0.01). (**C**) As in (A)/(B) but average VCN/cell; (Mean + SEM, n=4; **p<0.01).

3.6.1 Functional assay of candidate UCOEs in undifferentiated P19 and F9 cells

The novel candidate UCOE vectors were able to confer partial stability of expression in undifferentiated P19 (Figure 3.4) and F9 (Figure 3.6) cells. We next evaluated the ability of these novel LVs to stabilize expression upon differentiation of P19 and F9 cells down the neuroectodermal and endodermal lineages, respectively.

Cells transduced with each LV were first induced to form embryoid bodies in retinoic acid-containing differentiation medium via culture on non-adhesive plastic, which is a necessary pre-requisite state for later neuronal and endodermal differentiation. Cells subsequently differentiate into a mix of neuroectodermal and parietal endoderm cell types, of which fibroblast-like cells were the predominant type at the initial stage of differentiation (2-3 days after plating for P19 cells and 5-6 days for F9 cells), but with neurons and endoderm cells also present. The cultures were then enriched for non-dividing cells by the application of 10 μM cytosine arabinoside, in which neurons and endoderm cells appear as the only survivors. The optimum incubation time to form embryoid bodies and to differentiate the cells into neuroectodermal and endoderm cell types was determined as 2-3 days for P19 cells and 5-6 days for F9 cells.

Prior to differentiation, the quality of the P19 and F9 cell cultures and starting percentage of eGFP+ cells was determined by flow cytometry and shown to be acceptable (Figures 3.7 and 3.8).

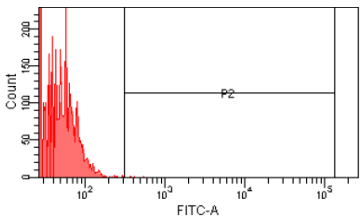
The expression results from triplicate samples of cells transduced with the 1.5A2UCOE-SEW, SET-CCN-SEW, CCN-SET-SEW, B1-RPS-SEW, RPS-B1-SEW LVs following differentiation is presented in Figures 3.8 and 3.10. Our data from the differentiation experiments show that expression from the control LV carrying the A2UCOE construct (1.5A2UCOE-SEW) has remained stable after the P19 and F9 cells were differentiated down the neuronal and endoderm lineages, respectively (Figures 3.8-A and 3.10-A). However, the novel candidate UCOE vectors SET-CCN-SEW, CCN-SET-SEW, B1-RPS-SEW and RPS-B1-SEW show a gradual reduction in eGFP expression whilst the silencing control SEW vector was rapidly repressed (Figures 3.8-A and 3.10-A).

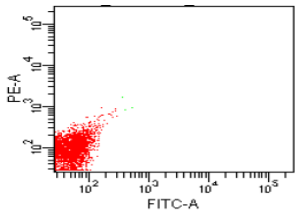
As in the case of undifferentiated cells (Figures 3.4 and 3.6), MFI values upon differentiation of both P19 and F9 cells paralleled the eGFP+ cell expression data (Figures 3.8-B and 3.10-B) whilst average vector copy number was essentially stable throughout the timecourse of the experiment (Figures 3.8-C and 3.10-C).

Thus the LVs tested in these experiments gave similar outcomes in undifferentiated and differentiated F9 and P19 cells. However, it is perhaps noteworthy that the rate of reduction in eGFP expression with the SEW and

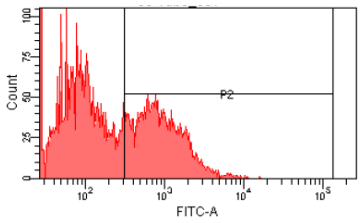
novel candidate UCOE LVs is greater upon differentiation (Figures 3.8-A and 3.10-A) than in undifferentiated (Figures 3.4-A and 3.6-A) cells.

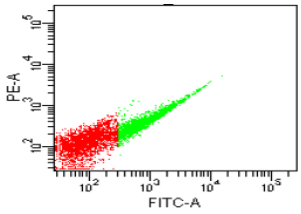
A: NEG



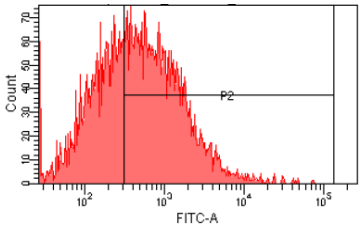


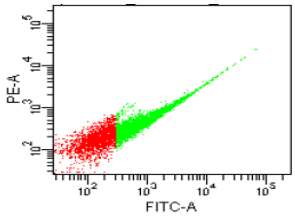
B: SEW



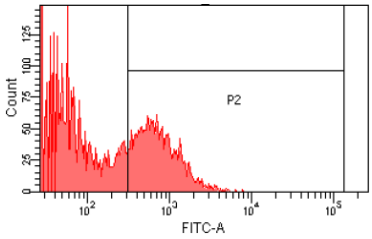


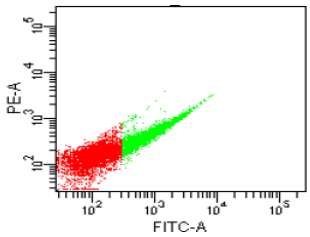
C: 1.5A2UCOE -SEW



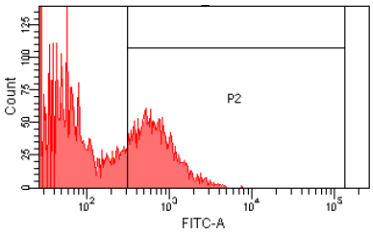


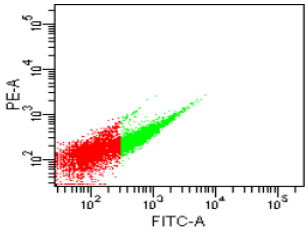
D: SET-CCN -SEW



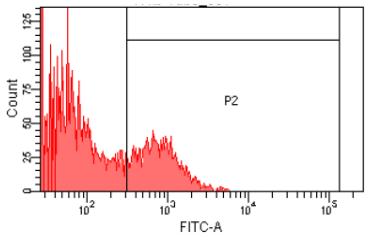


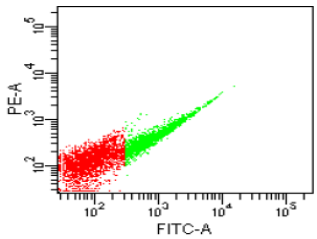
E: CCN-SET -SEW



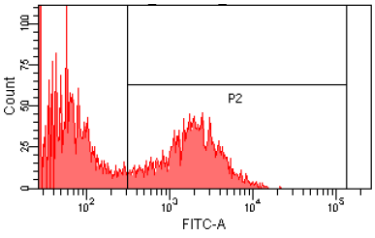


F: B1-RPS -SEW





G: RPS-B1 -SEW



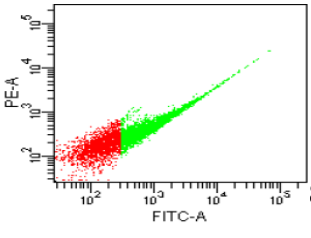
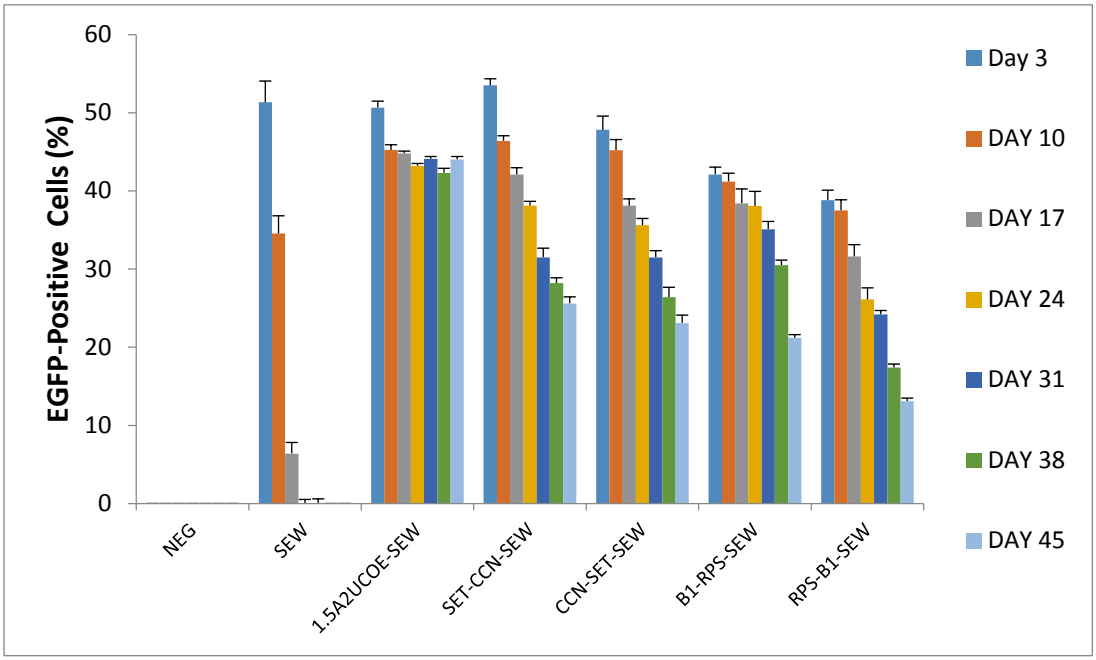
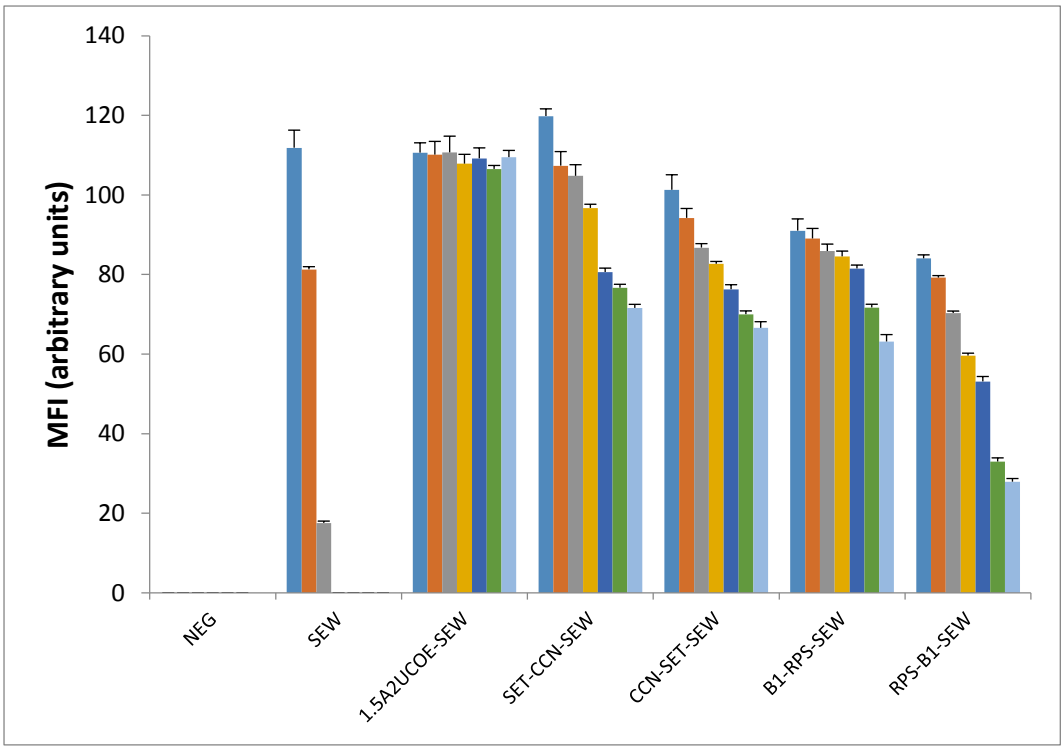


Figure 3.7: Flow cytometry plots of untransduced and transduced P19 cells prior to differentiation down the neuroectodermal lineage. P19 cells were transduced with the LVs at MOI 3 and cells analysed by flow cytometry 3 days later. GFP-negative cells are shown with red; GFP-positive cells are shown in green. (**A**) Negative control-untransduced cells. (**B**) SEW transduced cells. (**C**) 1.5A2UCOE-SEW transduced cells. (**D**) SET-CCN-SEW transduced cells (**E**) CCN-SET-SEW transduced cells (**F**) B1-RPS-SEW transduced cells (**G**) RPS-B1-SEW transduced cells.

A)


EGFP-Positive Cells (%)
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
SET-CCN-SEW
CCN-SET-SEW
B1-RPS-SEW
RPS-B1-SEW
Day 3
DAY 10
DAY 17
DAY 24
DAY 31
DAY 38
DAY 45

B)


MFI (arbitrary units)
140
120
100
80
60
40
20
0
NEG
SEW
1.5A2UCOE-SEW
SET-CCN-SEW
CCN-SET-SEW
B1-RPS-SEW
RPS-B1-SEW

C)

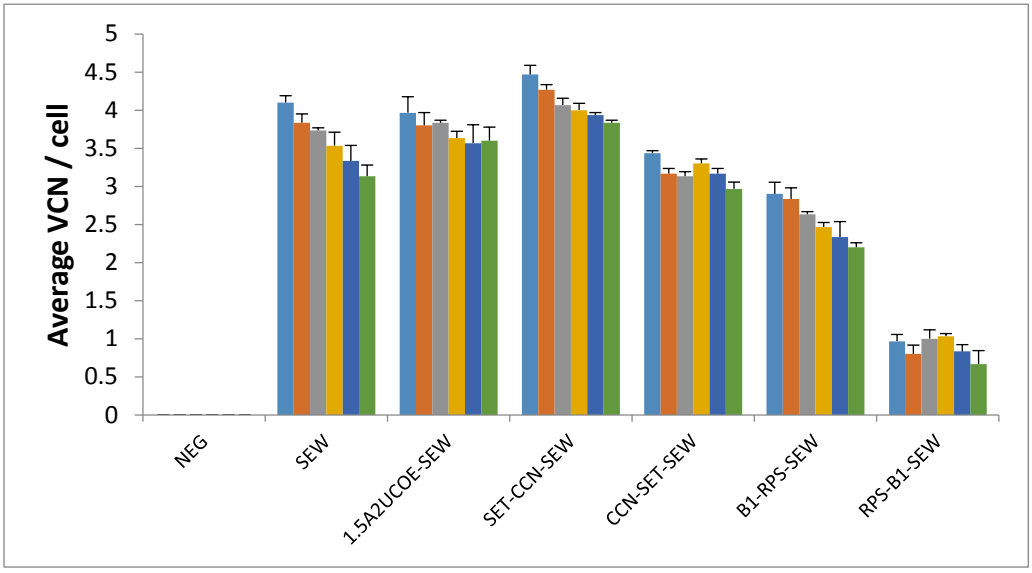
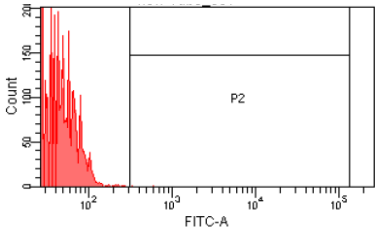
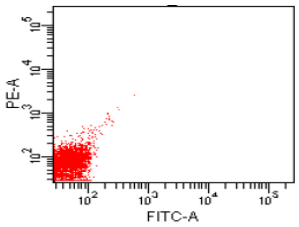


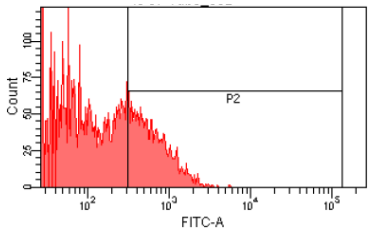
Figure 3.8: Novel candidate UCOEs offer only partial protection against silencing in differentiated P19 cells. P19 cells were transduced with novel candidate UCOE SET-CCN-SEW, CCN-SET-SEW, B1-RPS-SEW, RPS-B1-SEW and control SEW and 1.5A2UCOE-SEW LVs (Figure 3.1) and at 3-days post-transduction induced to undergo differentiation down the neuroectodermal lineage. Cells were analysed at various times following differentiation by flow cytometry to detect percentage of eGFP reporter gene expressing (eGFP+) cells, mean fluorescence intensity (MFI) and by real-time Q-PCR for average vector copy number (VCN) per cell. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 45 days post-transduction (**A**) timecourse of percentage eGFP positive cells; (Mean + SEM, n=4; **p<0.01). (**B**) As in (A) but showing MFI; (Mean + SEM, n=4; **p<0.01). (**C**) As in (A)/(B) but average VCN/cell; (Mean + SEM, n=4; **p<0.01).

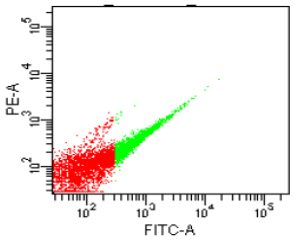
A: NEG



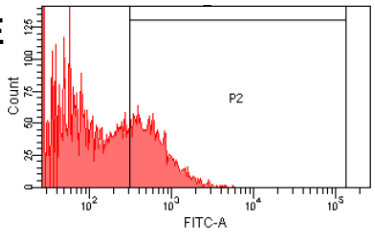


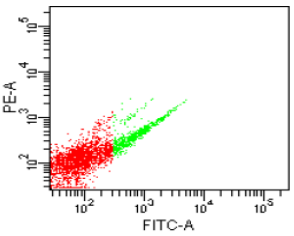
B: SEW



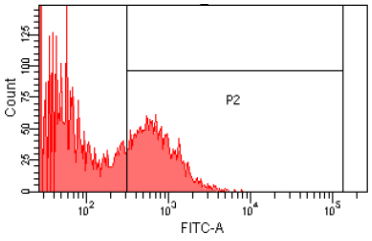


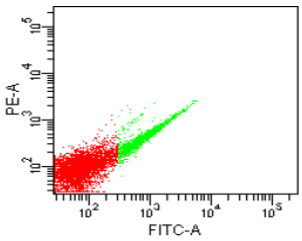
C: 1.5A2UCOE -SEW



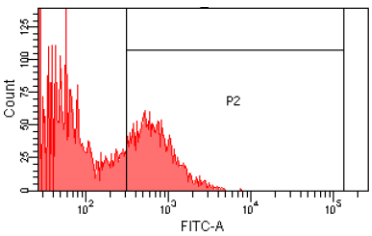


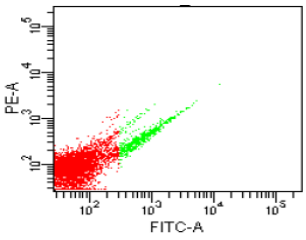
D: SET-CCN -SEW



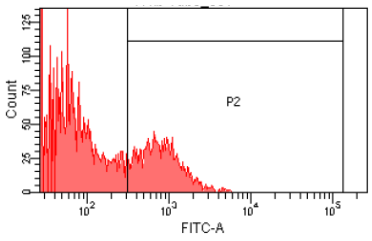


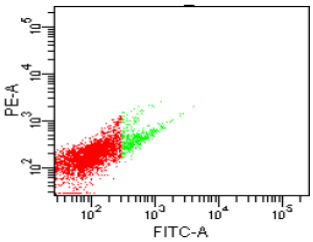
E: CCN-SET -SEW



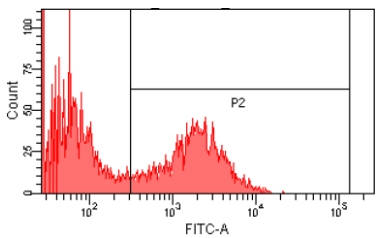


F: B1-RPS -SEW





G: RPS-B1 -SEW



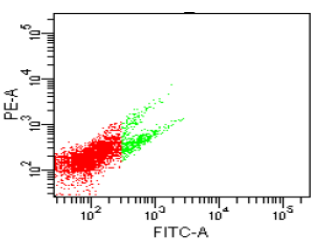
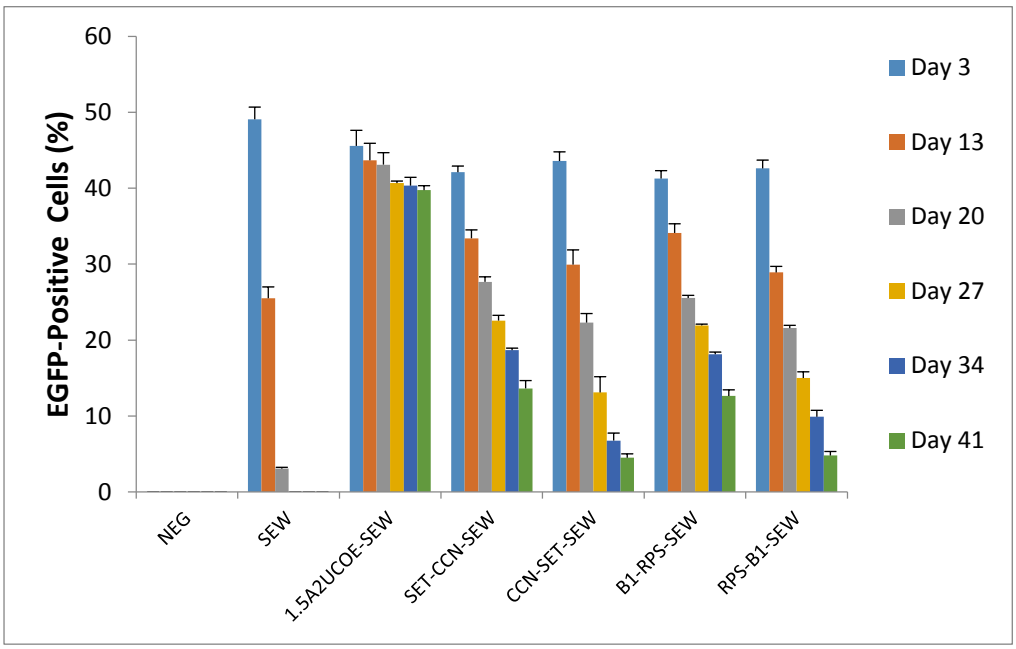
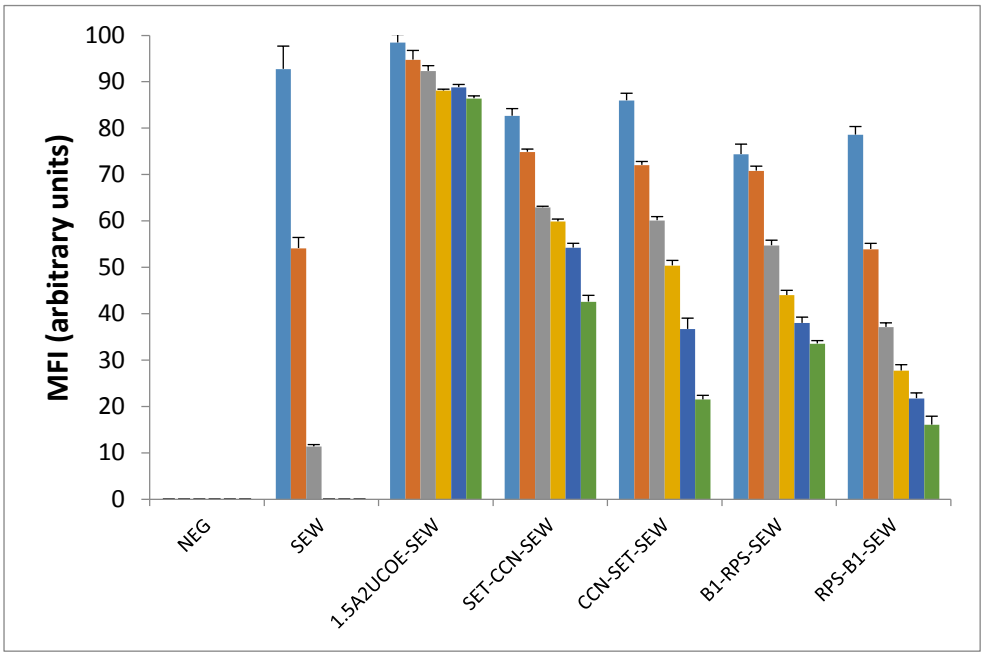


Figure 3.9: Flow cytometry plots of untransduced and transduced F9 cells prior to differentiation down the endodermal lineage. F9 cells were transduced with novel candidate UCOE SET-CCN-SEW, CCN-SET-SEW, B1-RPS-SEW, RPS-B1-SEW and control SEW and 1.5A2UCOE-SEW lentiviral vectors (Figure 3.1). Cells were analyzed by flow cytometry 3 days post-transduction. GFP-negative cells are shown in red; GFP-positive cells are shown in green. (**A**) Negative control-untransduced cells. (**B**) SEW transduced cells. (**C**) 1.5A2UCOE-SEW transduced cells. (**D**) SET-CCN-SEW transduced cells. (**E**) CCN-SET-SEW transduced cells. (**F**) B1-RPS-SEW transduced cells. (**G**) RPS-B1-SEW transduced cells.

A)


EGFP-Positive Cells (%)
60
50
40
30
20
10
0
Day 3
Day 13
Day 20
Day 27
Day 34
Day 41
NEG
SEW
1.5A2UCOE-SEW
SET-CCN-SEW
CCN-SET-SEW
B1-RPS-SEW
RPS-B1-SEW

B)


MFI (arbitrary units)
100
90
80
70
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
SET-CCN-SEW
CCN-SET-SEW
B1-RPS-SEW
RPS-B1-SEW

C)

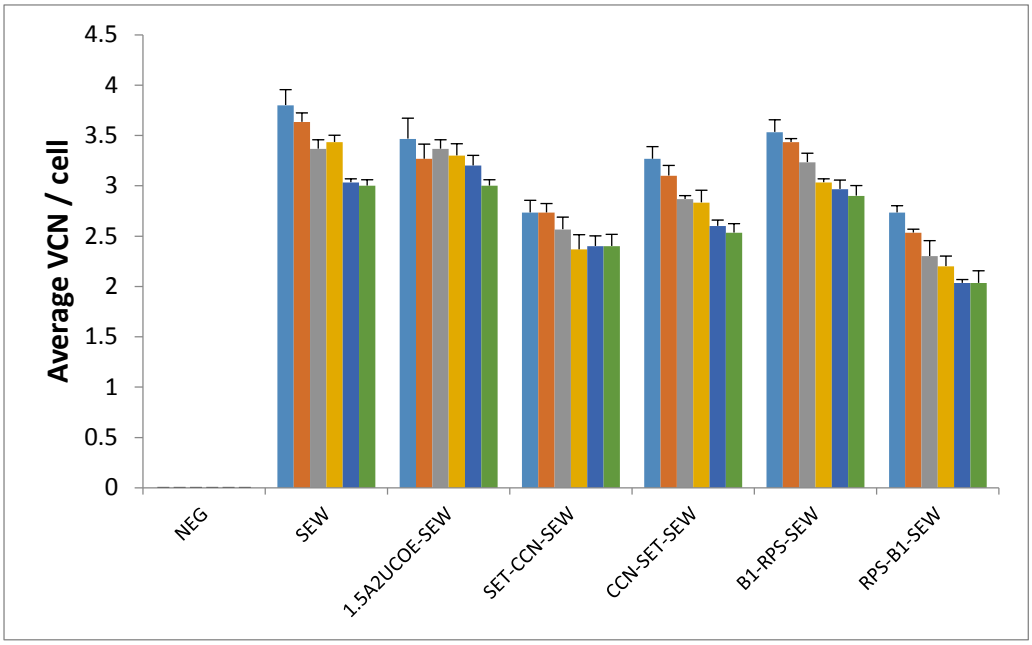
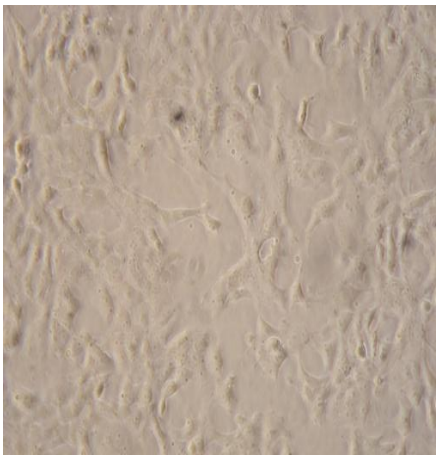


Figure 3.10: Novel candidate UCOEs offer only partial protection against silencing in differentiated F9 cells. F9 cells were transduced with novel candidate UCOE SET-CCN-SEW, CCN-SET-SEW, B1-RPS-SEW, RPS-B1-SEW and control SEW and 1.5A2UCOE-SEW LVs (Figure 3.1). At 3-days post transduction cells were induced to undergo differentiation down the endodermal lineage. Cells were analysed by flow cytometry at periodic intervals to detect percentage of eGFP reporter gene expressing (eGFP+) cells, mean fluorescence intensity (MFI) and by real-time Q-PCR for average vector copy number (VCN) per cell. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 41 days post-transduction. (A) timecourse of percentage eGFP positive cells; (Mean + SEM, n=4; **p<0.01). (B) As in (A) but showing MFI; (Mean + SEM, n=4; **p<0.01). (C) As in (A)/(B) but average VCN/cell; (Mean + SEM, n=4; **p<0.01). Note: 3-day results are prior to induced neuronal differentiation.

3.7 Immunofluorescence staining and quantification of transduced differentiated P19 and F9 cells

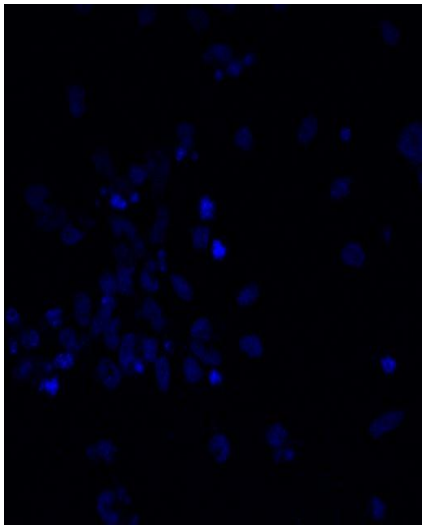
In order to confirm the results obtained by flow cytometry (Figures 3.9 and 3.10), P19 and F9 cells which had undergone neuroectodermal and endodermal differentiation, respectively, were stained for appropriate markers and scored under immunofluorescence microscopy. This was undertaken as follows: After differentiation of cells to embryoid bodies in uncoated bacterial dishes, cells were transferred into 6 well-plates containing protein-coated glass cover slips. Upon completion of the period of differentiation, cells were stained for neuroectodermal (P19 cells) and endodermal (F9 cells) differentiation markers and scored by immunofluorescence microscopy for single eGFP or differentiation marker positive cells or double positive eGFP plus differentiation marker positive cells. More precisely, P19 and F9 cells were scored for immunofluorescence staining following differentiation after plating on $1cm^2$ laminin-coated glass slides. For detection of neuroectodermal differentiation, P19 cells were immunofluorescence stained with anti-β-III tubulin (1:500) as primary and rhodamine-antimouse Alexa Fluoro 546 (1:200) as secondary antibody. Detection of endodermal differentiation, F9 cells were stained for Oct3-4 (1:200) as primary antibody and rabbit anti-mouse Alexa Fluoro 488 (1:500) as secondary antibody. Figure 3.11 shows representative microscopic images of undifferentiated and differentiated P19 and Figure 3.12 similarly for F9 cells.

A

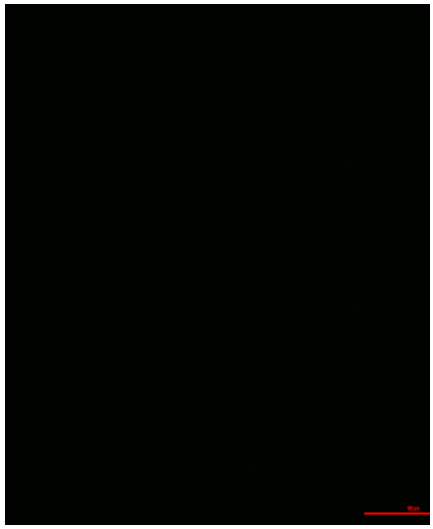


B

C

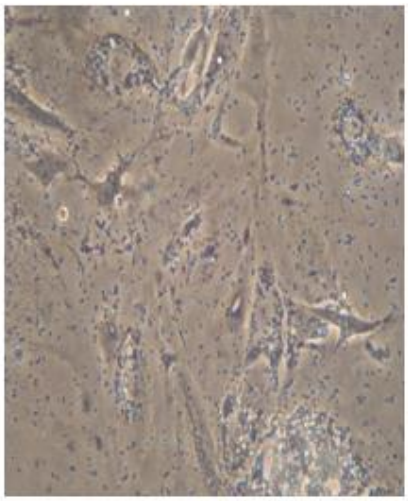


D



E

F



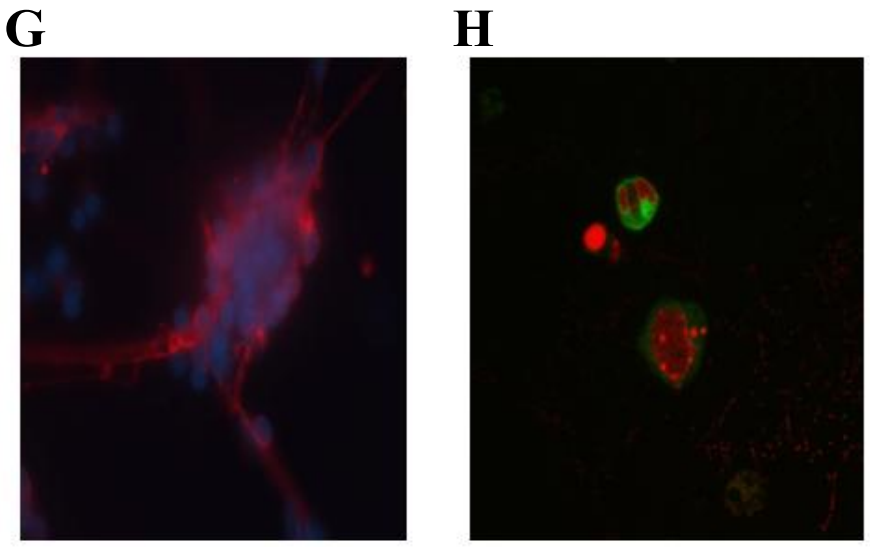
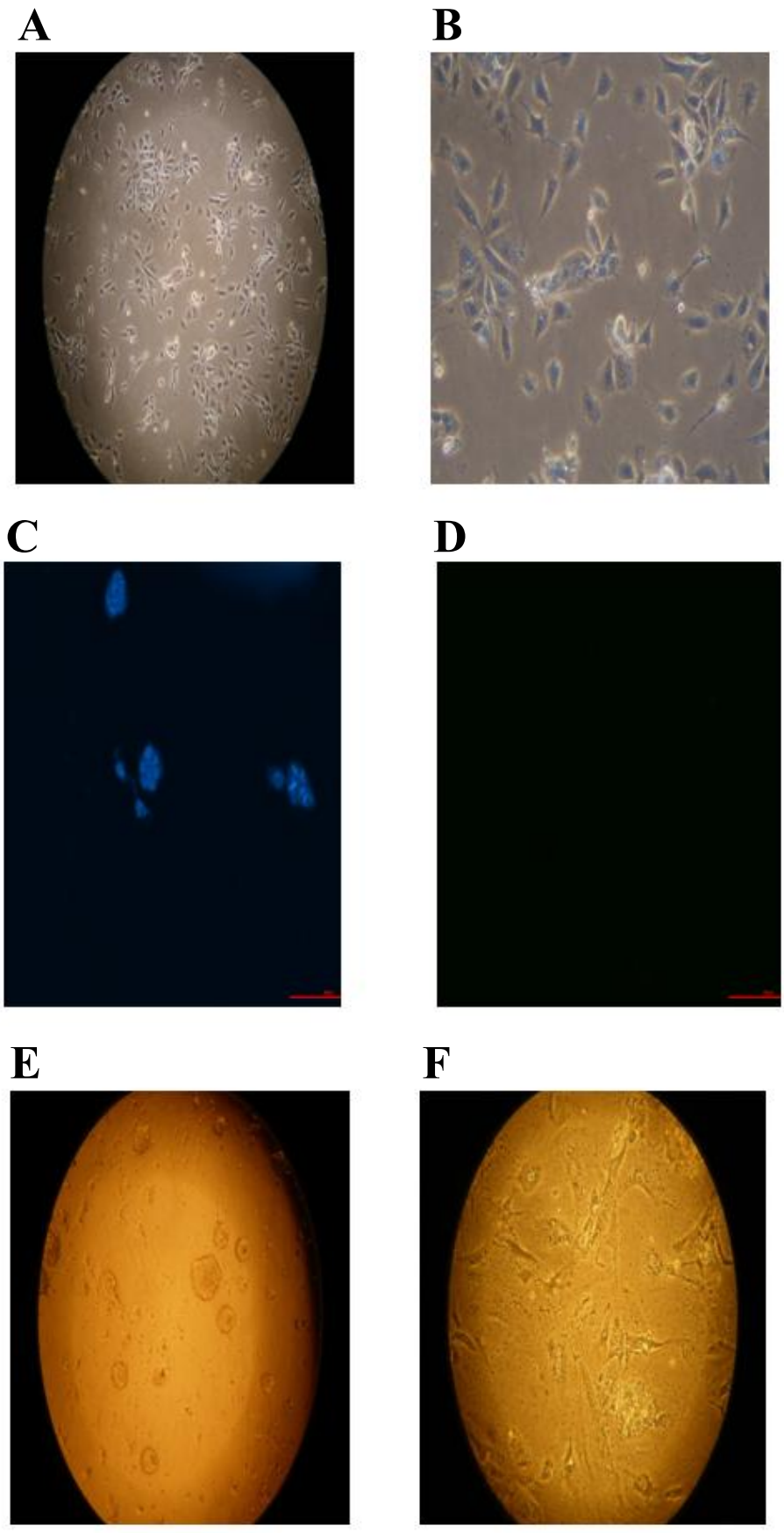


Figure 3.11. Microscopic images of undifferentiated and differentiated P19 cells.
A-B: Light microscopy of undifferentiated and untransduced P19 cells; 40X magnification.
C-D: Undifferentiated and untransduced P19 cells after immunofluorescence staining with DAPI (panel C) and anti-β-III tubulin (panel D); 40x and 100x magnification respectively. As expected cells are only positive for DAPI (blue) staining with neither anti-β-III tubulin (red) nor GFP (green) fluorescence.
E-F: Embryoid bodies (panel E) formed after 2 days of culture in uncoated bacterial dishes; 40x magnification. Panel F, P19 cells following 4 days of differentiation down the neuroectodermal lineage after embryoid body transfer to normal coated tissue culture plates; 100x magnification.
G-H: Immunofluorescence staining of neuroectodermal differentiated P19 cells transduced with a UCOE-eGFP vector; 100x magnification. Panel G, DAPI (blue) and anti-β-III tubulin (red) staining. Panel H, anti-β-III tubulin (red) and eGFP (green) fluorescence shown.


A
B
C
D
E
F

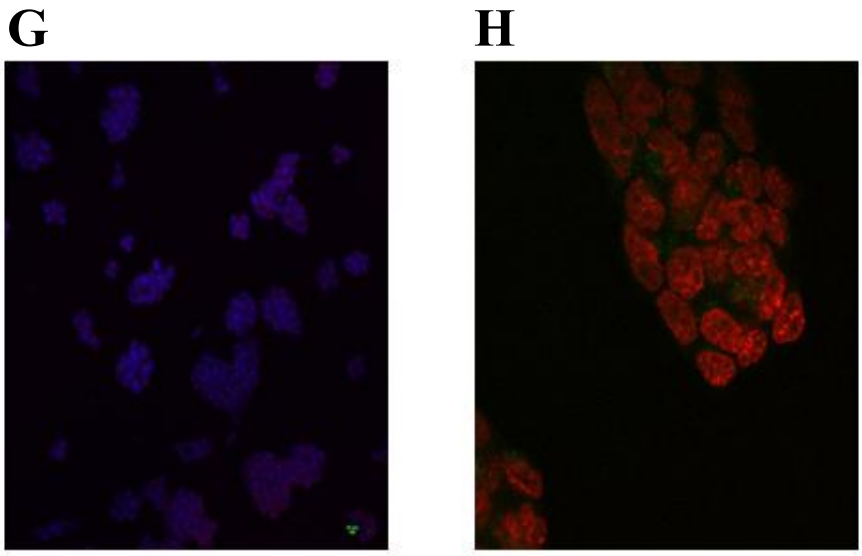


Figure 3.12: Microscopic images of undifferentiated and differentiated F9 cells.
A-B: Light microscopy of undifferentiated and untransduced F9 cells. Panel A, 20x magnification; panel B, 40x magnification.
C-D: Undifferentiated and untransduced F9 cells after immunofluorescence staining with DAPI (panel C) and anti-Oct3-4 (panel D); 40x magnification. As expected cells are only positive for DAPI (blue) staining with neither anti-Oct3-4 (red) nor GFP (green) fluorescence.
E-F: Embryoid bodies (panel E) formed after 2 days culture in uncoated bacterial dishes; 40x magnification. Panel F, F9 cells following 5 days of differentiation down the endodermal lineage after embryoid body transfer to normal coated tissue culture plates; 100x magnification.
G-H: Immunofluorescence staining of endodermal differentiated F9 cells transduced with a UCOE-eGFP vector. Panel G, DAPI (blue) and anti-Oct3-4 (red) staining; 20x magnification. Panel H, anti-Oct3-4 (red) and eGFP (green) fluorescence shown; 100x magnification.

Quantification of immunofluorescence stained cells was as follows: first, all cells stained with these markers on the slides were scored. Then, the eGFP positive cells that had been transduced with the UCOE-eGFP LVs were also counted in the same slides. Between 70 to 90 for P19, and 250 to 300 for F9 cells were counted in total on each slide to obtain the average number of differentiated cells at in 24 days post-differentiation induction. The results (Figures 3.13 and 3.14) show the average number of eGFP plus differentiation marker double positive cells and confirm those obtained by flow cytometry. Expression from the positive control 1.5A2UCOE-SEW vector remained stable over the 24 day experimental period whilst the negative control SEW LV was silenced within the first two weeks. All novel candidate UCOE vectors showed a partial stabilising effect.

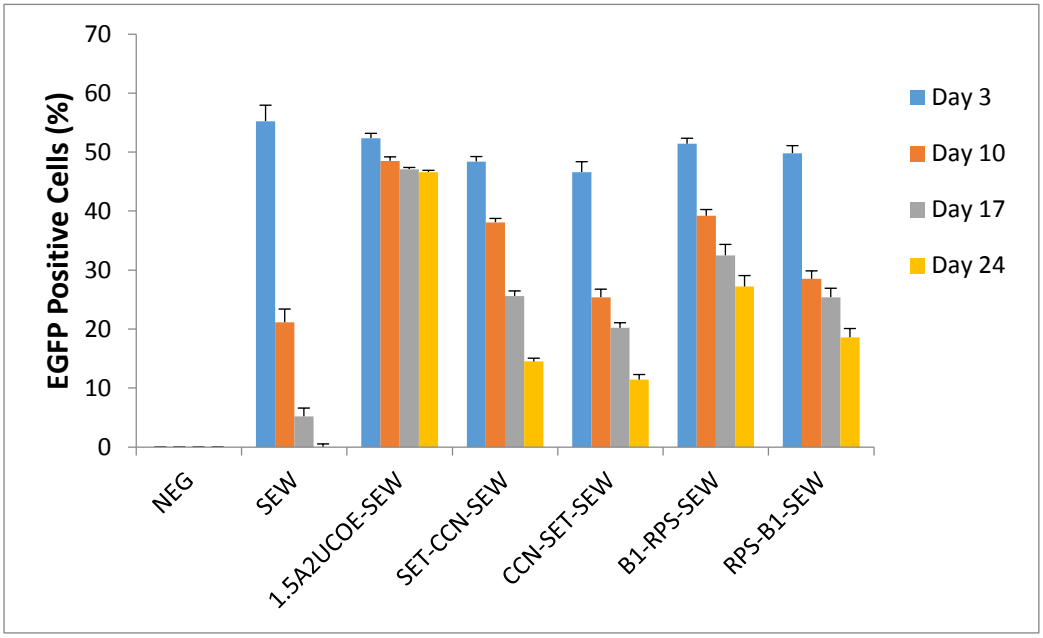


Figure 3.13: Novel candidate UCOEs offer only partial protection against silencing in differentiated P19 cells. P19 cells were transduced with novel candidate UCOE SET-CCN-SEW, CCN-SET-SEW, B1-RPS-SEW, RPS-B1-SEW and control SEW and 1.5A2UCOE-SEW LVs (Figure 3.1) and at 3-days post-transduction induced to undergo differentiation down the neuroectodermal lineage. Cells were analysed at various times following differentiation by fluorescence microscope by scoring for eGFP and anti-β-tubulin III double positive cells. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 24 days post-transduction (Mean + SEM, n=4; **p<0.01).

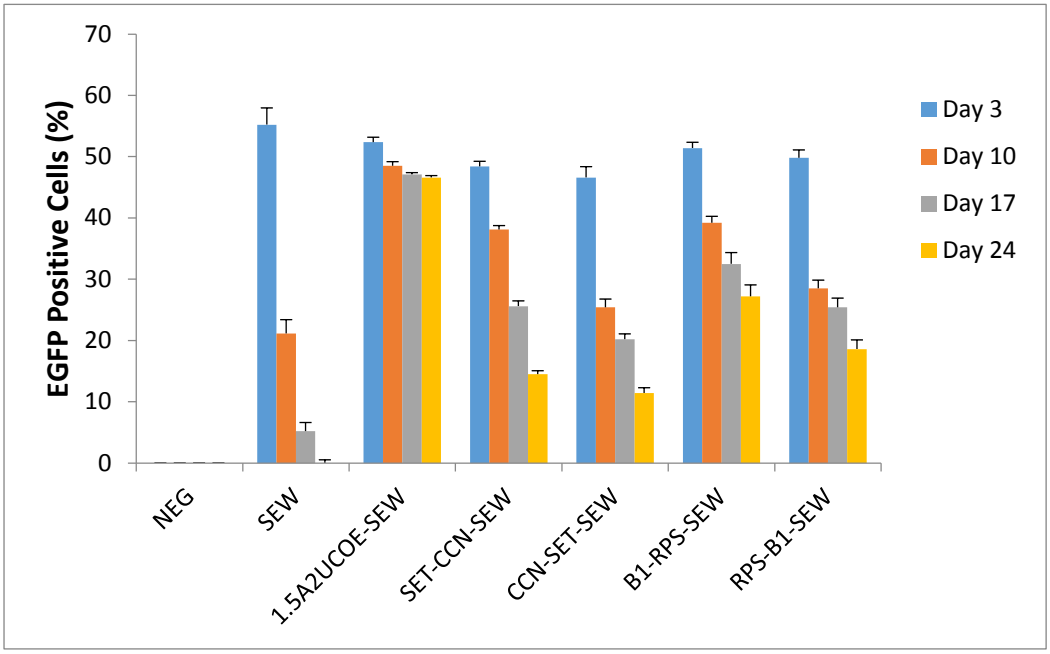


Figure 3.14: Novel candidate UCOEs offer only partial protection against silencing in differentiated F9 cells. F9 cells were transduced with novel candidate UCOE SET-CCN-SEW, CCN-SET-SEW, B1-RPS-SEW, RPS-B1-SEW and control SEW and 1.5A2UCOE-SEW LVs (Figure 3.1) and at 3-days post-transduction induced to undergo differentiation down the endodermal lineage. Cells were analysed at various times following differentiation by fluorescence microscope by scoring for eGFP and anti-Oct3-4 double positive cells. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 24 days post-transduction (Mean + SEM, n=4; **p<0.01).

3.8 Summary and Conclusions

The primary aim of the experiments conducted in this phase of the project was to determine whether the novel candidate UCOEs (Figure 3.1) conform with the dual divergent transcription model of A2UCOE function (Antoniou, Harland et al. 2003, Williams, Mustoe et al. 2005, Allen and Antoniou 2007) and are able to confer stability of expression on a linked heterologous promoter regardless of their orientation. Although the A2UCOE can confer stability of expression on a linked ubiquitous (Williams, Mustoe et al. 2005, Zhang, Frost et al. 2010) or tissue-specific (Talbot, Waddington et al. 2009, Brendel, Müller-Kuller et al. 2011) promoter it does so in an orientation specific manner with the CBX3 end of the element abutting the linked promoter. Our starting hypothesis was that this problem could be circumvented if both promoters of the selected gene pairs showed similar levels and variance of expression in a variety of tissues. That is, the reason why uniformity of expression levels both across tissues but also in between the gene pair was taken into consideration. In this regard our results are encouraging in that both orientations of the SETD3-CCNK element (SET-CCN-SEW, CCN-SET-SEW) and the artificial HNRPA2B1-RPS11 combination (B1-RPS-SEW, RPS-B1-SEW) showed similar abilities to confer stability of expression from the SFFV promoter independent of orientation (Figures 3.4 and 3.8).

Stability levels of expression of the novel candidate UCOE vectors as well as the controls was analysed in both undifferentiated and neuroectodermal and endodermal differentiated P19 and F9 cells, respectively. Our study has thus

extended previous investigations, which only looked at expression of UCOE-based LV profiles in undifferentiated P19 cells (Yoon, Lee et al. 2009, Zhang, Frost et al. 2010). The results provide a comparison of the previously reported results of the A2UCOE vectors in P19 cells as well as that of the control vectors. It is also clear from our results that neither the SETD3-CCNK nor the HNRPA2B1-RPS11 elements were as effective as the 1.5A2UCOE-SEW vector at negating silencing of the linked SFFV promoter, with eGFP expression from the RPS-B1-SEW LV decreasing most rapidly. The 1.5A2UCOE-SEW gave perfectly stable expression over the course of our experiment in not only undifferentiated P19 cells as previously reported (Zhang, Frost et al. 2010), but also in F9 cells. Furthermore, 1.5A2UCOE-SEW vector maintained stable expression upon neuroectodermal and endodermal differentiation of P19 and F9 cells, respectively. Overall our results provide more evidence of the capacity of the A2UCOE present in the configuration found in 1.5A2UCOE-SEW to provide stability and reproducibility of transgene expression when linked to a heterologous promoter such as SFFV.

Given that the novel candidate UCOEs tested in this phase of the project did not match the efficacy of the original A2UCOE, it was decided to focus future work on further dissecting the A2UCOE in an effort to identify sub-regions that are particular crucial for its function.

Chapter 4
Results

Functional dissection of the *CBX3-HNRPA2B1* ubiquitous chromatin opening element (A2UCOE)

Chapter 4

Functional dissection of the CBX3-HNRPA2B1 ubiquitous chromatin opening element (A2UCOE)

4.1 Introduction and background

In this part of the study we investigated as to whether sub-regions of the *HNRPA2B1-CBX3* UCOE (A2UCOE) retained UCOE activity. This was prompted by a number of reports some of which call into question the need for associated promoter activity for UCOE function. The relative positions of the sub-regions of the A2UCOE tested are depicted in Figure 4.1.

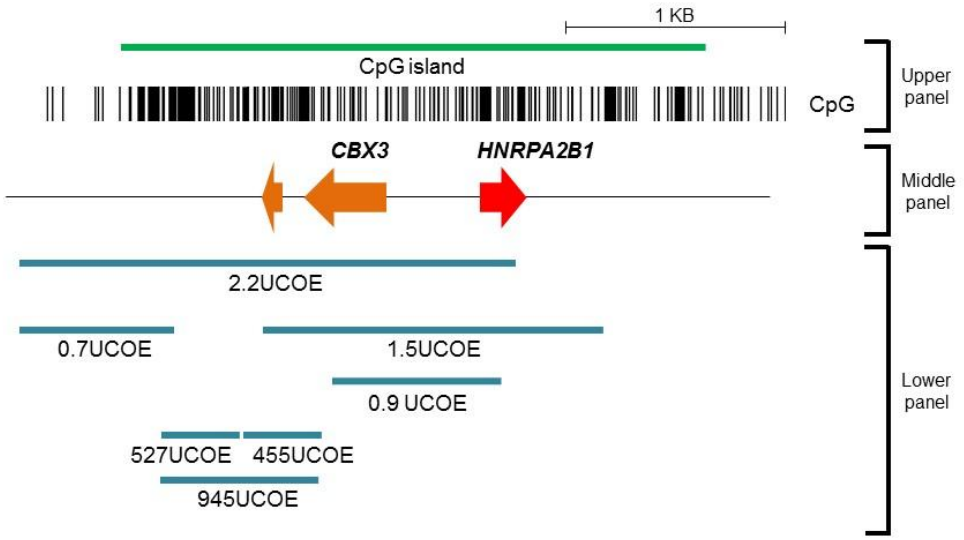


Figure 4.1. Illustration of the relative positions of the A2UCOE region sub-fragments analysed for UCOE activity. Upper panel: CpG density map and methylation-free CpG island. Middle panel: Positions of the two alternative first exons of *CBX3* (orange arrows) and the first exon of *HNRPA2B1* (red arrow). Lower panel: positions of A2UCOE sub-region fragments analysed for UCOE activity. 2.2UCOE: fully functional 2.2kb positive control element used to drive expression from *HNRPA2B1* promoter. 1.5UCOE: fully functional 1.5kb positive control element when linked upstream of the SFFV promoter. 0.7UCOE: "Daedalus" test fragment. 0.9UCOE: core test fragment predicted to constitute a methylation determining region (MDR). 945/527/455UCOE: test fragments spanning the most CpG dense region of the A2UCOE. The 0.7UCOE, 0.9UCOE and 945/527/455UCOE candidates would be tested for their ability to stabilise expression from the silencing prone SFFV promoter and compared against the positive control 1.5UCOE fragment.

The first was the publication by Bandarayake and colleagues, which claimed that a 700bp fragment from the 3' end of intron I of *CBX3* (Figure 4.2, 0.7UCOE), and thus devoid of promoter activity, was able to confer stability of expression on a linked CMV-GFP reporter construct from within an LV context in CHO cells (Bandaranayake, Correnti et al. 2011). If true these data would call into question the requirement for associated promoter activity for UCOE function. However, close scrutiny of the data presented clearly shows that although the percentage of GFP positive cells remains stable over time there is nevertheless a marked (55%) reduction in mean fluorescence intensity (MFI) suggesting transgene silencing at 4 weeks post-transduction (Figure 4.2B). The apparent discrepancy between stability in percentage of GFP-positive cells and instability in MFI appears to be due to high (8-10) average vector copy number as in principle only one of the integrated LVs need remain functional for a given cell to score positive in GFP expression. Based on these observations we wished to repeat these experiments but at a low average vector copy number per cell to accurately assess the stabilising capability of the promoter-less *CBX3* first intron derived 0.7UCOE element. We thus linked the same UCOE0.7, designated by these authors as "Daedalus", upstream of the SFFV promoter in forward and reverse orientations within the SEW LV (Figure 3.1) to generate the constructs Daedalus-F (Forward orientation 0.7kb UCOE-SFFV-eGFP) and Daedalus-R (Reverse orientation 0.7kb UCOE-SFFV-eGFP) (Figure 4.3).

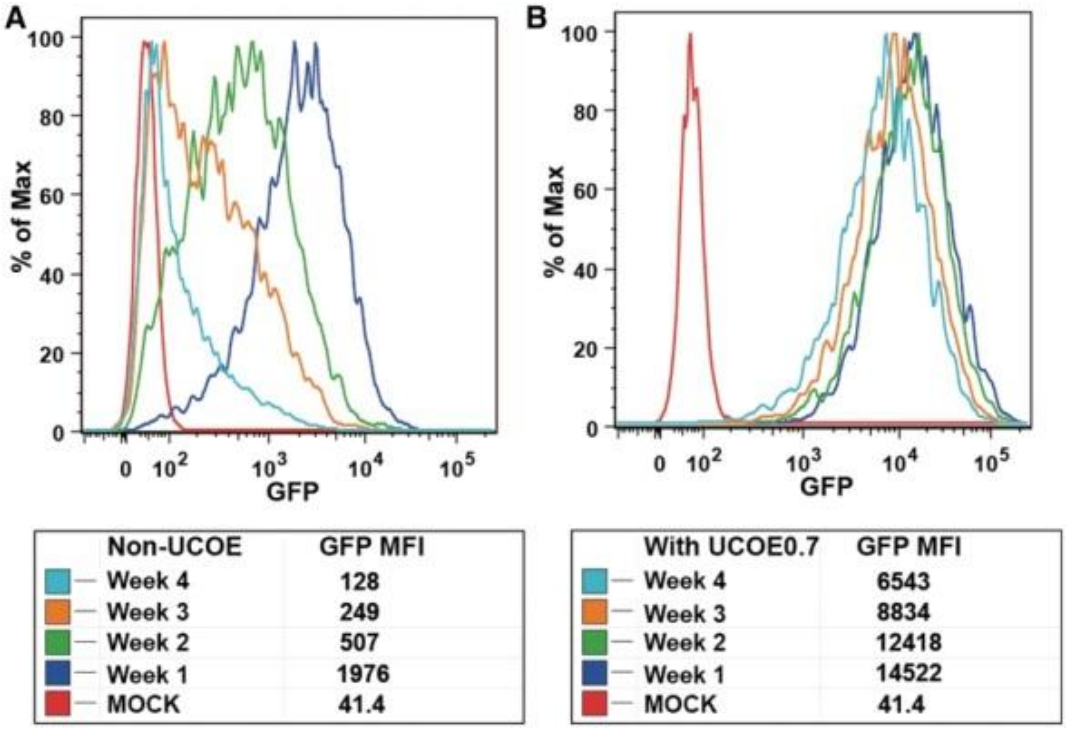


Figure 4.2. Lentiviral vectors containing either a CMV-GFP cassette alone or with 5' linked UCOE0.7 from the first intron of *CBX3* were used to transduce CHO cells and expression looked at over time by flow cytometry. (**A**) Results from the CMV-GFP construct show instability in both percentage of GFP-positive cells (upper panel) and MFI (lower table). (**B**) Results from UCOE0.7-CMV-GFP show stability in percentage GFP-positive cells (upper panel) but instability in MFI (lower table). Figure adapted from Bandarayake *et al*., 2011.

The second report was that of Thomson and colleagues (Thomson, Skene et al. 2010). This study shows data that suggest that a transgene consisting of a CpG-rich DNA fragment lacking a promoter is sufficient to establish a methylation-free region with associated active histone modification marks in the absence of RNA polymerase II. We, therefore, wished to analyse the most CpG dinucleotide-dense sub-regions of the A2UCOE, especially those lacking either the *HNRPA2B1* or *CBX3* promoters, for UCOE function. To this end, the region immediately downstream of the two alternative first exons of *CBX3* meet these criteria and were thus chosen for further functional analysis (Figure 4.1, 945UCOE, 527UCOE, 455UCOE). This 945bp region

was divided into two sub-fragments of 455bp and 527bp and all three linked individually upstream of the SFFV promoter within the SEW vector (Figure 4.3).

Lastly, a study by Lienert and colleagues shows that sub-fragments of CpG islands associated with the promoters of developmentally regulated genes that retain their ability to maintain appropriate DNA methylation status in a transgene context (Lienert, Wirbelauer et al. 2011). Some of the CpG island sub-fragments tested lacked their cognate promoter region. Due to this property, these CpG island sub-regions were designated as methylation-determining regions (MDRs). In order to ascertain if a similar MDR was present in the housekeeping gene region of the A2UCOE, a bioinformatics investigation was conducted to locate a region with specific transcription factor binding sites (Sp1, CTCF, USF) that constitute such an element. This resulted in a core region of the A2UCOE extending from the first exons of *CBX3* and *HNRPA2B1* and thus including both promoter regions being a possible MDR (data not shown). In order to test this hypothesis two constructs were built where a 0.9kb core A2UCOE region encompassing the hypothetical MDR (Figure 4.1) was linked upstream of the SFFV promoter within the SEW LV to test its ability to confer stability of expression on the latter (Figure 4.3; 0.9UCOE-F and 0.9UCOE-R).

These vectors were functionally analysed for UCOE activity by transducing murine embryonal carcinoma P19 and F9 cells, our proven system for evaluating stability of viral vector expression (Chapter 3; Zhang *et al.*, 2007 and 2010). The stability, potency, and reproducibility of expression of the eGFP transgene were then analysed in differentiated and undifferentiated P19 and F9 cells.

4.2. Lentiviral vectors used in this study

Note: the LV constructs Daedalus-F, Daedalus-R, 0.9UCOE-F, and 0.9UCOE-R LVs were constructed by Kristian Skipper. The bioinformatics analysis leading to the location of a putative MDR element within the A2UCOE, and subsequently incorporated into the 0.9UCOE-F and 0.9UCOE-R, was also conducted by Kristian Skipper. The 945UCOE, 527UCOE, 455UCOE vectors were constructed by our collaborators Fang Zhang and Adrian Thrasher at the Institute of Child Health (UCL, UK). All are depicted in Figure 4.3.

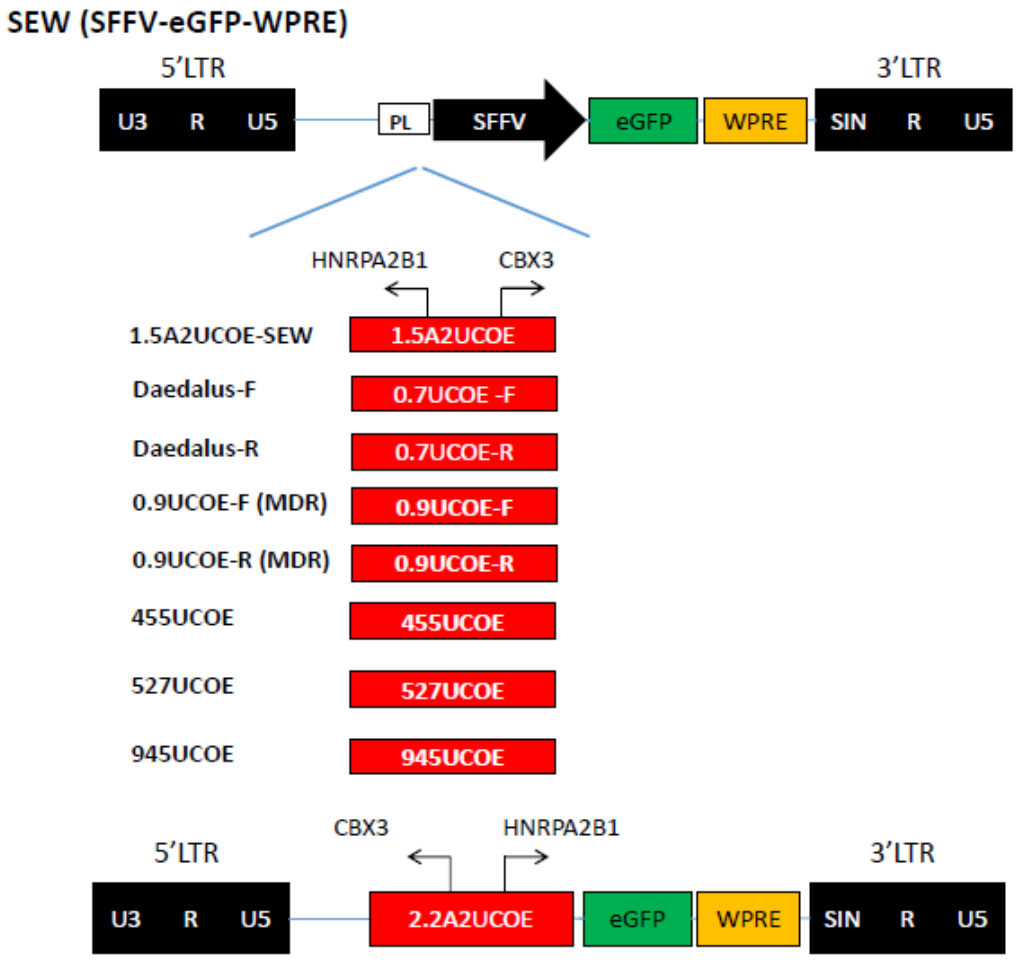


Figure 4.3. Illustration of the novel candidate UCOE and control lentiviral vectors. A standard self-inactivating (SIN) lentiviral vector backbone containing a spleen focus forming virus (SFFV) promoter driving expression of an enhanced green fluorescent protein (eGFP) reporter gene with downstream woodchuck hepatitis virus post-transcriptional regulatory element (WPRE) designated as SEW acted as the starting point for the construction of test UCOE vectors. The positive control 1.5kb *HNRPA2B1-CBX3* UCOE (1.5A2UCOE) was inserted into a polylinker (PL) cloning site upstream of the SFFV promoter to give 1.5A2UCOE-SEW. The second positive control 2.2kb *HNRPA2B1-CBX3* UCOE (2.2A2UCOE) is linked directly to the GFP reporter gene driving expression off the *HNRPA2B1* promoter to give the construct designated 2.2A2UCOE. Novel candidate UCOEs: 0.7UCOE 0.7kb Daedalus-F (forward) and Daedalus-R (reverse) constructs; 0.9kb putative methylation determining region (MDR) test constructs 0.9UCOE-F (forward), 0.9UCOE-R (reverse) orientations; CpG rich region candidate UCOE sub-fragments of 455bp 455(UCOE), 527bp 527UCOE and 945bp (945UCOE). Novel candidate UCOEs were inserted into the PL upstream of the SFFV promoter in SEW. LTR: long terminal repeat.

4.2.1 Daedalus-F and Daedalus-R

Daedalus-F (0.7UCOE-SFFV-eGFP) and Daedalus-R (reverse orientation; 0.7UCOE-R-SFFV-eGFP) constructs were derived from the *CBX3* first intron (Figure 4.1, lower panel) linked in both orientations upstream of the SFFV promoter (Figure 4.3). This 0.7kb candidate UCOE is as previously described (Bandaranayake, Correnti et al. 2011).

4.2.2 0.9UCOE-F and 0.9UCOE-R

The 0.9kb central core A2UCOE region (Figure 4.1, lower panel) constituting a hypothetical MDR (methylation determining region; (Lienert, Wirbelauer et al. 2011)) linked upstream of the SFFV promoter in forward (F) and reverse (R) orientations within SEW (Figure 4.3).

4.2.3 455UCOE, 527UCOE, 945UCOE

Fragments encompassing the most CpG dense regions within intron I of *CBX3* of 455bp, 527bp and 945bp in length (Figure 4.1, lower panel) linked upstream of the SFFV promoter in their native orientation within SEW (Figure 4.3). These were obtained from our collaborators at the Institute of Child Health (UCL, UK). They are based on a reported finding that promoter-less CpG rich regions can establish an open chromatin structure (Thomson, Skene et al. 2010). Thus, CpG rich regions from the CBX3 side of the A2UCOE and devoid of promoter activity were tested to see if they

could retain chromatin-opening capability and confer stability of expression from a linked SFFV promoter.

4.3 Lentiviral vector titration in HEK293T cells

Cells were collected for analysis by flow cytometry three days after the transduction of $2x10^5$ HEK293T cells with serial dilutions of harvests at both 2 and 3 days post-transfection from each production run for all LVs: SEW (SFFV-eGFP-WPRE), 1.5A2UCOE-SEW, 2.2A2UCOE, Daedalus-F/R, 0.9UCOE-F/R, 455UCOE, 527UCOE and 945UCOE (Table 4.1).

LV	Volume (µL) of viral vector stock			
	2 µL	0.2 µL	0.02 µL	0.002 µL
	Percentage eGFP+ cells 1st / 2nd viral harvest			
SEW	84.01 / 71.22	15.28 / 11.52	3.11 / 1.86	0.6 / 0.42
1.5A2UCOE-SEW	74.43 / 51.3	11.3 / 9.14	1.8 / 1.44	0.7 / 0.4
2.2A2UCOE	83.4 / 74.66	18.26 / 13.62	2.24 / 1.83	0.5 / 0.3
DAEDALUS-F	64.98/62.18	9.5 / 8.4	1.2 / 1.01	0.2 / 0.1
DAEDALUS-R	76.21 / 62.14	7.25 / 6.32	1.44 / 1.11	0.22 / 0.11
0.9UCOE-F	86.6 / 75.2	8.66 / 6.55	2.32 / 1.87	0.00 / 0.00
0.9UCOE-R	63.1 / 55.4	5.26 / 3.15	0.82 / 0.57	0.00 / 0.00
455UCOE	76.1 / 72.2	13.76 / 11.45	2.86 / 1.88	0.4 / 0.2
527UCOE	83.66 / 72.22	18.46 / 13.15	3.44 / 2.17	0.6 / 0.4
945UCOE	63.2 / 52.4	8.15 / 5.35	1.32 / 0.77	0.00 / 0.00

Table 4.1. Titration of lentiviral vector (LV) preparations. A $2x10^5$ aliquot of HEK293T cells were transduced with 0.002-2 µl of a given LV stock from the first and second harvest of virus from culture supernatant during the time of preparation. Analysis of eGFP positive cells at 3-days post-transduction was by flow cytometry. LV constructs are as described in Figure 2.1.

4.3.1 Calculation of LV titre

Viral titre (Table 4.2) was calculated based on the dilution of each harvest that gave a percentage of eGFP-positive cells between 1% and 10%. Pools of cells with percentages of eGFP-positive cells higher than 10% are very likely to contain multiple integrations of the vector, while very low scores could be false positives.

On day one, $2x10^5$ 293T cells were seeded in each 6-well. Thus, the percentage of eGFP-positive cells from flow cytometry analysis, reflects the percentage of the initial cell population that was successfully transduced and hence the number of infectious units added in the specified well. The dilution factor or the volume of the lentivirus preparation that was used is known, so the calculation of the number of infectious units per ml (iv/mL) from that point onwards is straightforward.

LV	Titre (iu/mL)	
	1st harvest	2nd harvest
SEW	2.88 x 10^8	1.23x 10^8
1.5A2UCOE-SEW	1.46 x 10^8	1.14 x 10^8
2.2A2UCOE	2.17 x 10^8	1.84 x 10^8
DAEDALUS-F	1.51 x 10^8	1.05 x 10^8
DAEDALUS-R	1.82 x 10^8	1.12 x 10^8
0.9UCOE-F	1.26 x 10^8	0.63 x 10^8
0.9UCOE-R	1.73 x 10^8	1.03 x 10^8
455UCOE	1.51 x 10^8	0.91 x 10^8
527UCOE	1.94 x 10^8	1.31 x 10^8
945UCOE	1.75 x 10^8	0.93 x 10^8

Table 4.2. Lentiviral vector (LV) titres. Viral titre as infectious units per ml (iu/mL) was calculated based on the dilution of 1st and 2nd harvest of virus that gave a percentage of eGFP-positive HEK293T cells between 1% and 10% (Table 4.1). LV constructs are as described in Figure 4.3.

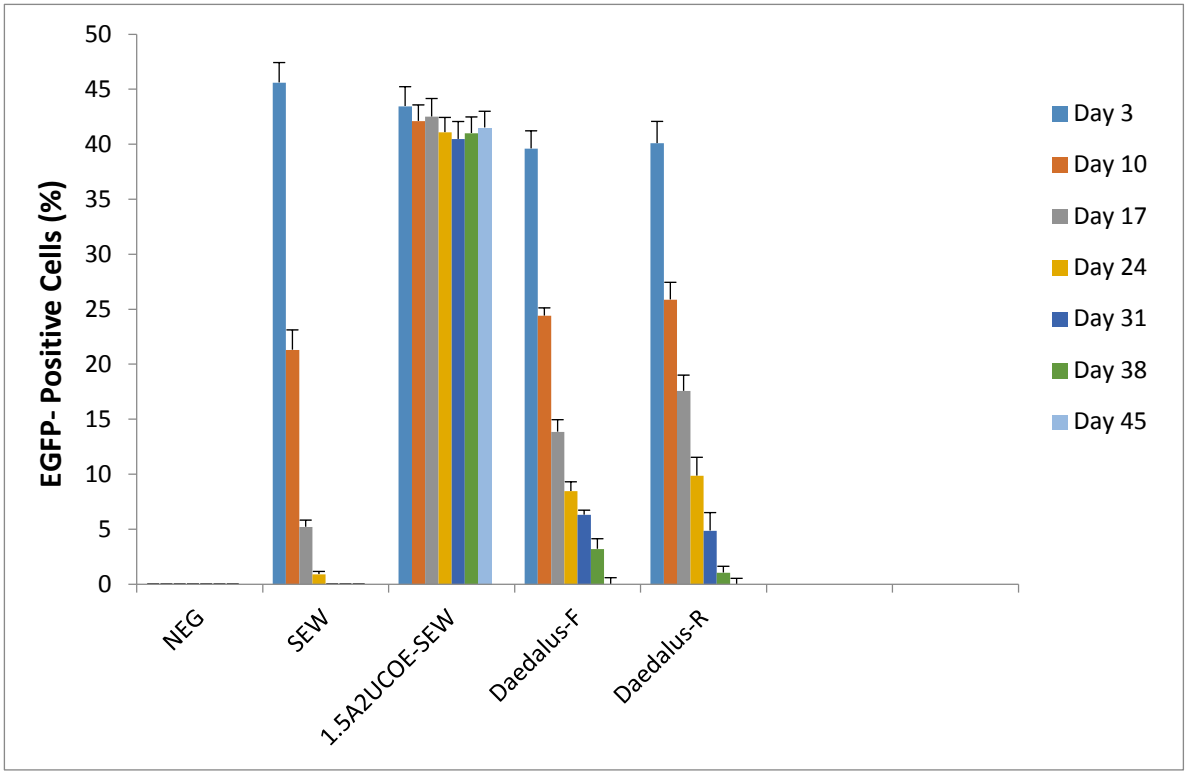
4.4 Functional analysis Daedalus-F and Daedalus-R candidate UCOE LVs in undifferentiated P19 and F9 cells

P19 and F9 cells were transduced with the generated LVs at a MOI of 3 or 6, with the intention to start the experiment within the range of 40-60% eGFP-positive cells in all pools. The transduced cells and negative control were then propagated in culture and assayed as a time course by flow cytometry, to determine the percentage of eGFP-positive cells and the mean fluorescence intensity (MFI) for each repeat transduction for each vector. Cell cultures were analysed every 7 days from day 3 post-transduction and extending to 45 days (Figure 4.4 and 4.5). Cell samples were taken at regular intervals for the flow cytometry analysis, to monitor the expression levels of the eGFP reporter gene. In addition, DNA was extracted from cells at each time point for RT-qPCR in order to determine the average vector copy number per cell.

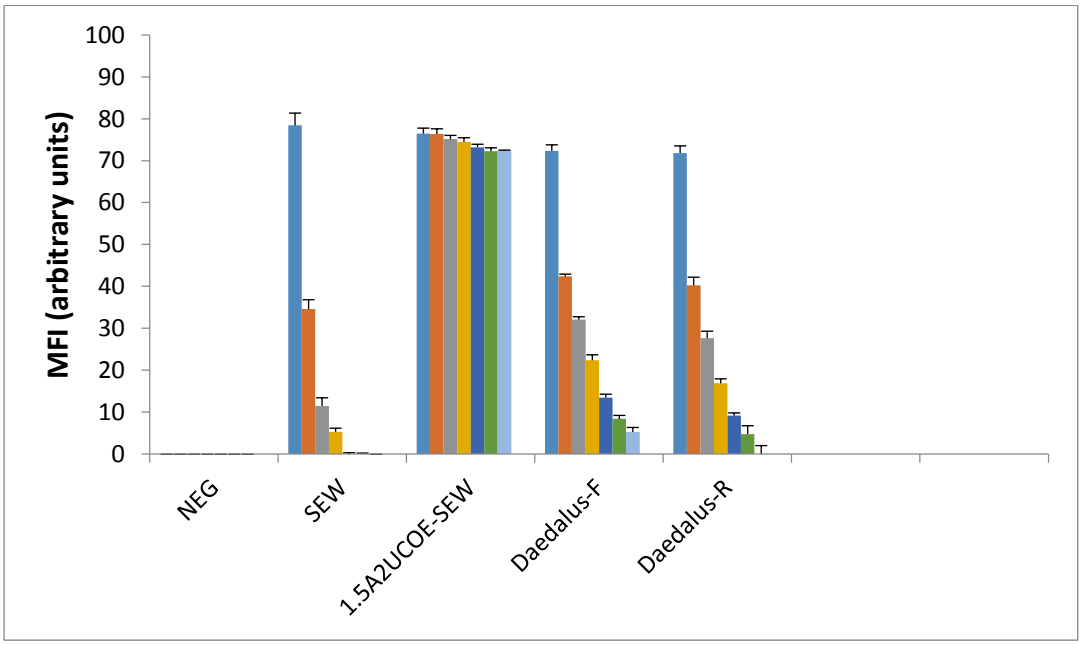
Figure 4.4A shows the flow cytometry time course results depicting percentage of eGFP+ cells. Although similar initial transduction efficiency was obtained with all vectors (45%-60% eGFP+ cells), expression from the SEW (SFFV-eGFP-WPRE), rapidly declined from 46 to 3% positive cells within 17 days. In contrast, the proportion of eGFP+ cells from the 1.5A2UCOE-SEW vector remained completely stable over the 45-day period of culture. Therefore, whilst differences in the percentage of eGFP-positive cells transduced with 1.5A2UCOE-SEW was significantly higher when compared with the results obtained for all other vectors, the decrease of eGFP expression of both the Daedalus-F and Daedalus-R vectors is parallel to that seen with SEW. The expression of eGFP in SEW transduced cells had dropped almost 80% at 2 weeks post-transduction. Similarly, with the exception of 1.5A2UCOE-SEW, the reduction observed in eGFP-positive cells was 40% to 50%

for all other vectors carrying the test Daedalus-F and Daedalus-R over the same time period. The values of mean fluorescent intensity (MFI) for all vectors paralleled the expression results on a percentage eGFP-positive cell basis (Figure 4.4B). MFI was stable in the 1.5A2UCOE-SEW transduced cells and unstable in the case of SEW and in the novel test Daedalus-F and Daedalus-R constructs.

A)


EGFP- Positive Cells (%)
50
45
40
35
30
25
20
15
10
5
0
NEG
SEW
1.5A2UCOE-SEW
Daedalus-F
Daedalus-R
Day 3
Day 10
Day 17
Day 24
Day 31
Day 38
Day 45

B)


MFI (arbitrary units)
100
90
80
70
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
Daedalus-F
Daedalus-R

C)

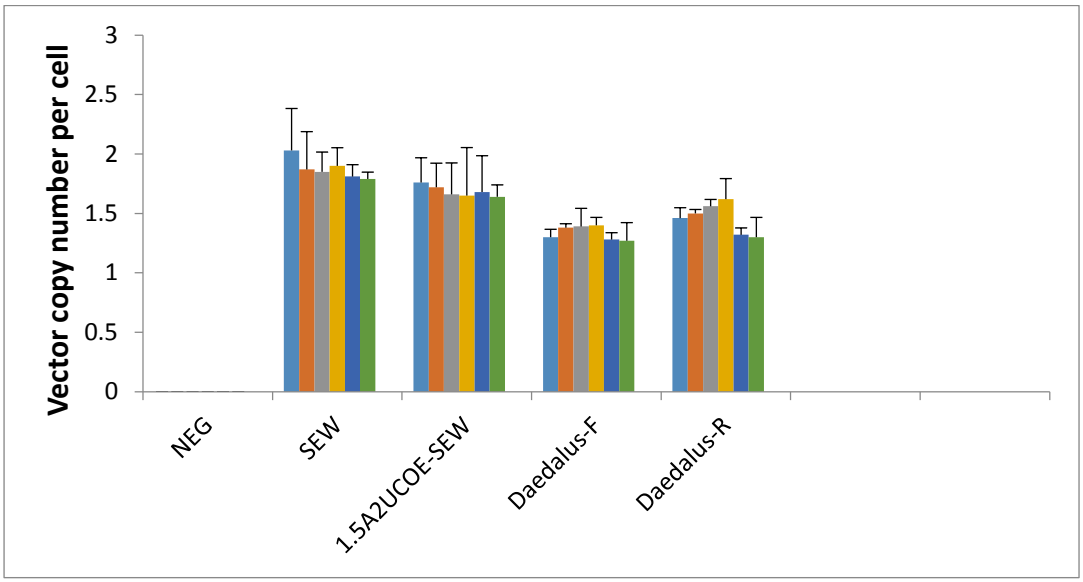
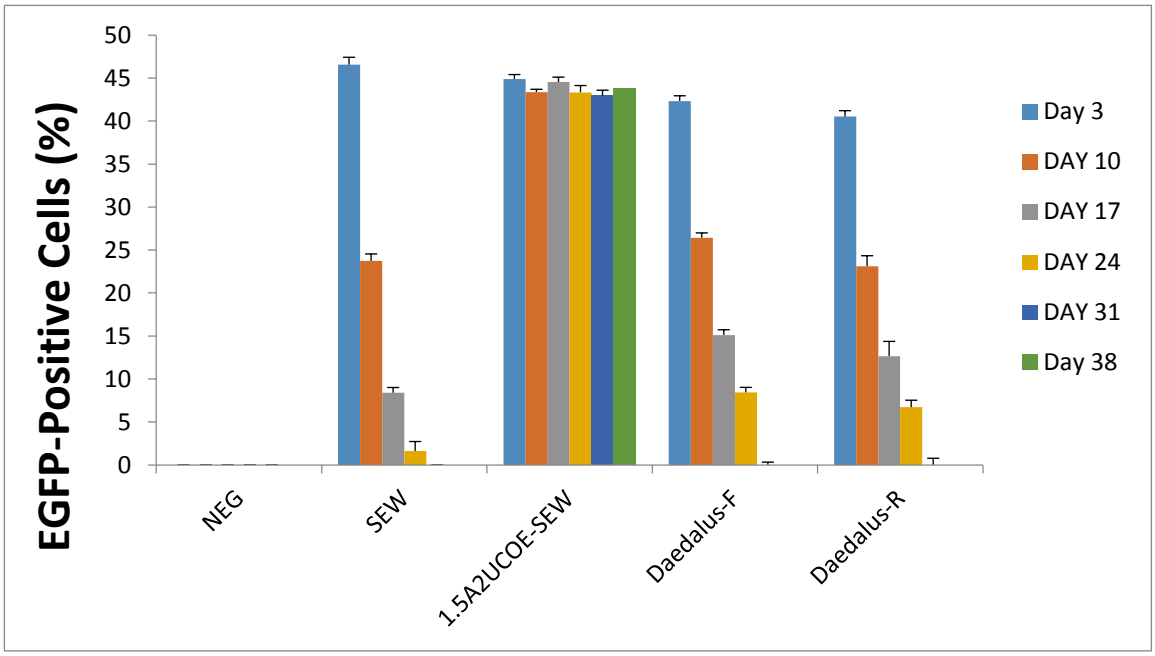


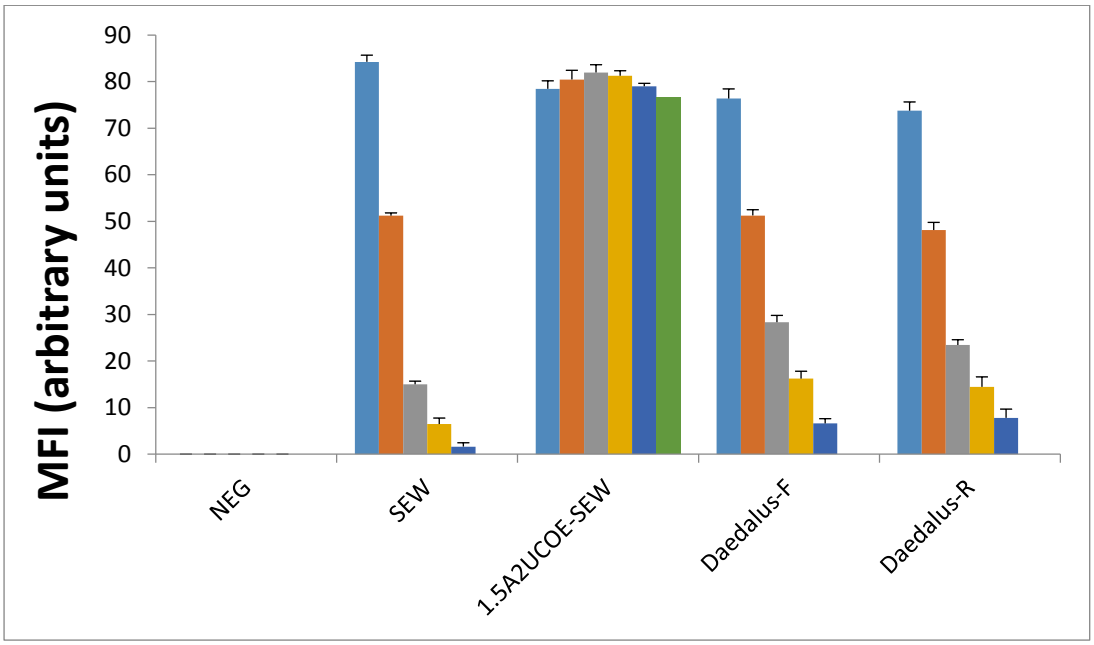
Figure 4.4. Novel candidate Daedalus-F and Daedalus-R UCOEs offer only partial protection against silencing in undifferentiated P19 cells. P19 cells were transduced with novel candidate UCOE Daedalus-F, Daedalus-R and control SEW and 1.5A2UCOE-SEW lentiviral vectors (Figure 4.3). Cells were analysed by flow cytometry to detect percentage of eGFP reporter gene expressing (eGFP+) cells, mean fluorescence intensity (MFI) and by RT-Q-PCR for average vector copy number (VCN) per cell. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 45 days post-transduction. (**A**) timecourse of percentage eGFP-positive cells; (Mean + SEM, n=4; **p<0.01). (**B**) As in (A) but showing MFI; (Mean + SEM, n=4; **p<0.01). (**C**) As in (A)/(B) but average VCN/cell; (Mean + SEM, n=4; **p<0.01).

Figure 4.5A shows the flow cytometry time course results depicting percentage of eGFP+ cells in transduced undifferentiated F9 cells. The results parallel those we observed in P19 cells (Figure 4.4A). Although a similar initial transduction efficiency was obtained with all vectors (39%-50% eGFP+ cells), expression from the SEW (SFFV-eGFP-WPRE), rapidly declined from 47% to 7% eGFP+ cells within 17 days. In contrast, the proportion of eGFP+ cells (~45%) from the 1.5A2UCOE-SEW vector remained completely stable over the 38-day period of culture. The UCOE candidate Daedalus-F and Daedalus-R constructs showed a very similar rate of decline in expression to that seen with SEW. The expression of eGFP in SEW transduced cells had dropped almost 80% after 2 weeks post-transduction. Similarly, with the exception of 1.5A2UCOE-SEW, the reduction observed in eGFP+ cells were 40% to 50% for the test UCOE Daedalus-F and Daedalus-R LVs over the same time period. The values of mean fluorescent intensity (MFI) for all vectors paralleled the eGFP expression results (Figure 4.5B). MFI was stable in the 1.5A2UCOE-SEW transduced cells, unstable in the case of SEW and partially stable in the novel test UCOE constructs.

A)


EGFP-Positive Cells (%)
50
45
40
35
30
25
20
15
10
5
0
NEG
SEW
1.5A2UCOE-SEW
Daedalus-F
Daedalus-R
Day 3
DAY 10
DAY 17
DAY 24
DAY 31
Day 38

B)


MFI (arbitrary units)
90
80
70
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
Daedalus-F
Daedalus-R

C)

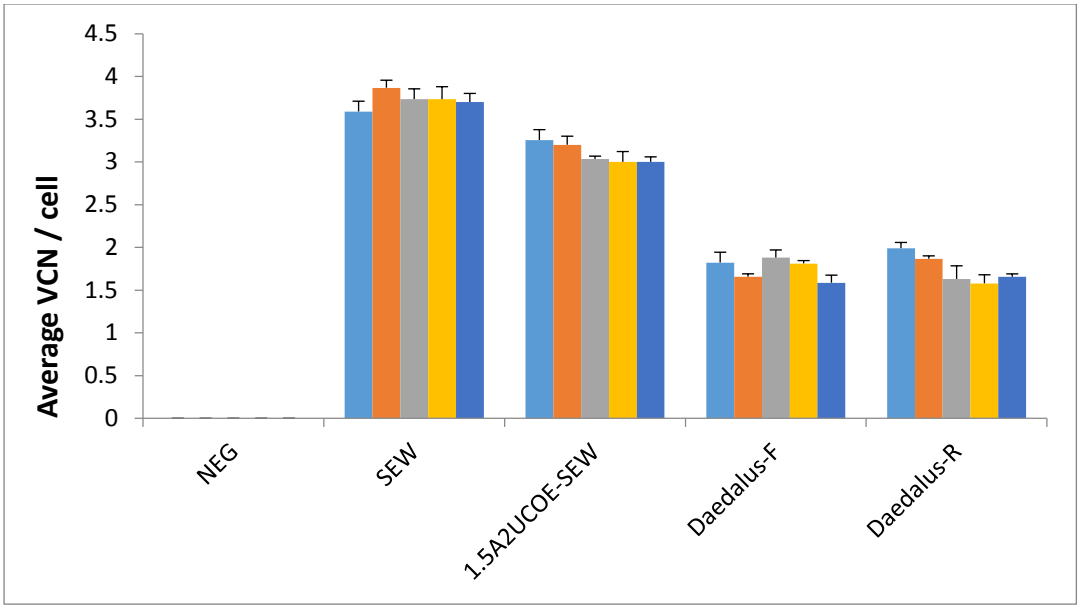


Figure 4.5. Novel candidate UCOE Daedalus-F and Daedalus-R offer only partial protection against silencing in undifferentiated F9 cells. F9 cells were transduced with novel candidate UCOE Daedalus-F and Daedalus-R and control SEW and 1.5A2UCOE-SEW LVs (Figure 4.3). Cells were analysed by flow cytometry to detect percentage of eGFP reporter gene expressing (eGFP+) cells, mean fluorescence intensity (MFI) and by real-time Q-PCR for average vector copy number (VCN) per cell. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 38 days post-transduction. (**A**) time course of percentage eGFP positive cells; (Mean + SEM, n=4; **p<0.01). (**B**) As in (A) but showing MFI; (Mean + SEM, n=4; **p<0.01). (**C**) As in (A)/(B) but average VCN/cell; (Mean + SEM, n=4; **p<0.01).

Average vector copy number per cell in both P19 (Figure 4.4C) and F9 (Figure 4.5C) cell transduction experiments remained stable throughout the period of culture demonstrating that reduction in eGFP-positive cells in the SEW and Daedalus-F and Daedalus-R constructs was not due to vector loss but to silencing of integration events.

4.5 Functional assay of candidate UCOEs Daedalus-F and Daedalus-R in differentiated P19 and F9 cells

The novel candidate UCOE Daedalus-F and Daedalus-R vectors were not able to confer stability of expression in undifferentiated P19 (Figure 4.4) and F9 (Figure 4.5) cells. We next evaluated the ability of these novel LVs to stabilize expression upon differentiation of P19 and F9 cells down the neuroectodermal and endodermal lineages, respectively.

Cells transduced with each LV were first induced to form embryoid bodies in retinoic acid-containing differentiation medium via culture on non-adhesive plastic, which is a necessary pre-requisite state for latter neuronal and endodermal differentiation. Cells subsequently differentiate into a mix of neuroectodermal and parietal endoderm cell types, of which fibroblast-like cells were the predominant type at the initial stage of differentiation (2-3 days after plating for P19 cells and 5-6 days for F9 cells), but with neurons and endoderm cells also present. The cultures were then enriched for non-dividing cells by the application of 10 μM cytosine arabinoside, in which neurons and

endoderm cells appear as the only survivors. The optimum incubation time to form embryoid bodies and to differentiate the cells into neuroectodermal and endoderm cell types was determined as 2-3 days for P19 cells and 5-6 days for F9 cells.

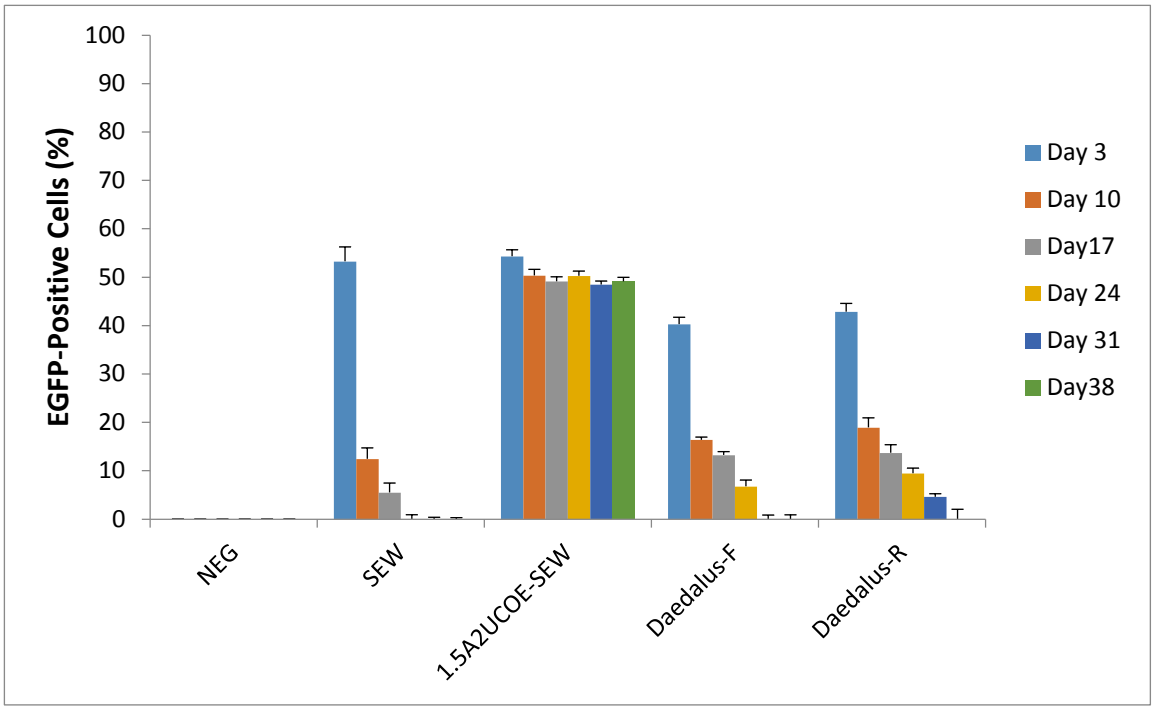
The expression results obtained by flow cytometry of the total cell population from triplicate samples of cells transduced with the SEW, 1.5A2UCOE-SEW, Daedalus-F and Daedalus-R LVs following differentiation is presented in Figures 4.6 and 4.7. Our data from the differentiation experiments show that the positive control LV carrying the 1.5kb A2UCOE construct (1.5A2UCOE-SEW) has remained stable after the P19 and F9 cells were differentiated down the neuronal and endoderm lineages, respectively (Figures 4.6A and 4.7A). However, the novel candidate UCOE vectors Daedalus-F and Daedalus-R show a rapid reduction in eGFP expression whilst eGFP expression in the silencing control SEW vector was also rapidly repressed (Figures 4.6A and 4.7A).

As in the case of undifferentiated cells (Figures 4.4B and 4.5B), MFI values upon differentiation of both P19 and F9 cells paralleled the eGFP+ cell expression data (Figures 4.6B and 4.7B) whilst average vector copy number was essentially stable throughout the timecourse of the experiment (Figures 4.6C and 4.7C).

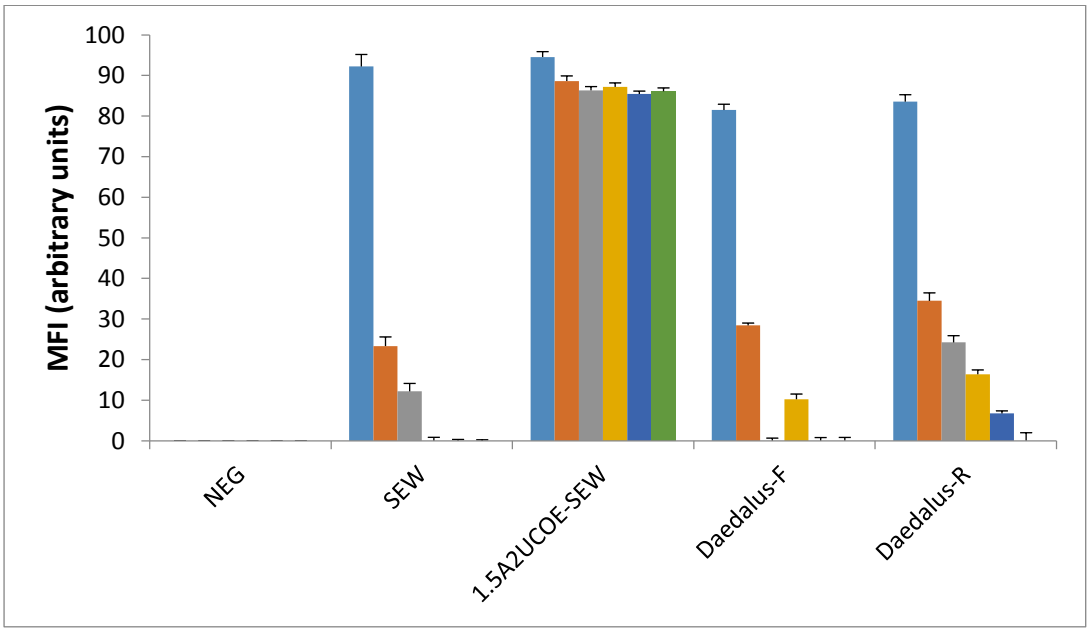
Thus, the LVs tested in these experiments gave similar outcomes in undifferentiated and differentiated F9 and P19 cells. However, it is perhaps noteworthy that the rate of reduction in eGFP expression with the SEW and novel candidate UCOE LVs is greater upon differentiation (Figures 4.6A and 4.7A) than in undifferentiated (Figures 4.4A and 4.5A) cells.

As the Daedalus-F and Daedalus-R vectors did not provide stable expression in differentiated P19 or F9 cells, it was deemed unnecessary to conduct immunofluorescence staining for neuroectodermal and endodermal markers.

A)


EGFP-Positive Cells (%)
100
90
80
70
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
Daedalus-F
Daedalus-R
Day 3
Day 10
Day17
Day 24
Day 31
Day38

B)


MFI (arbitrary units)
100
90
80
70
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
Daedalus-F
Daedalus-R

C)

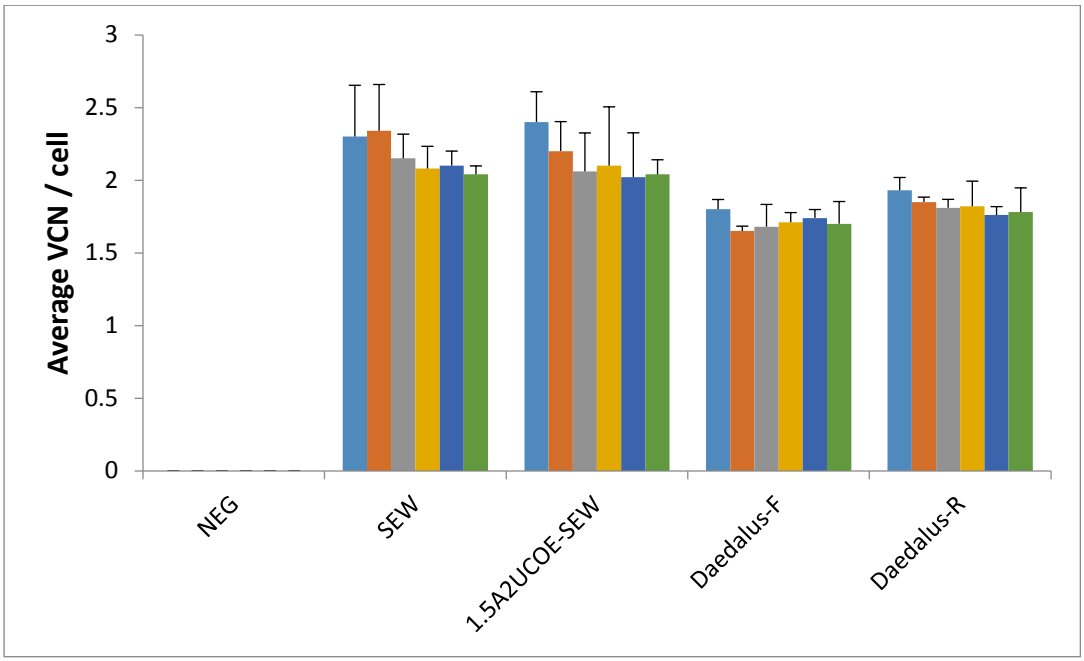
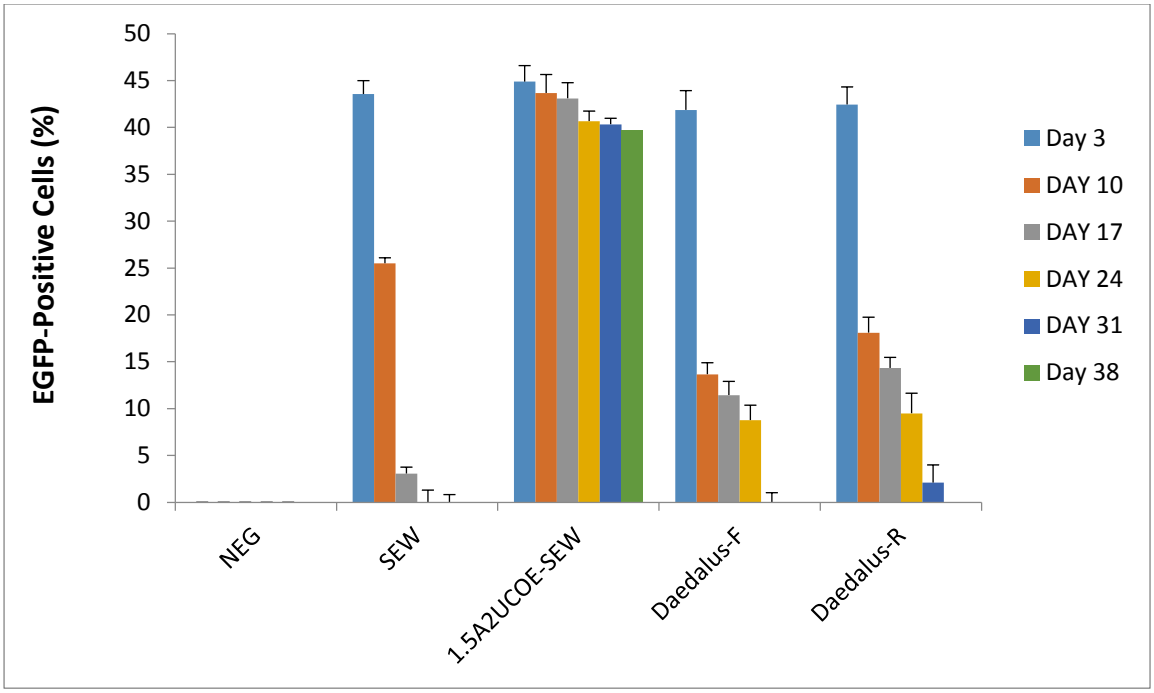
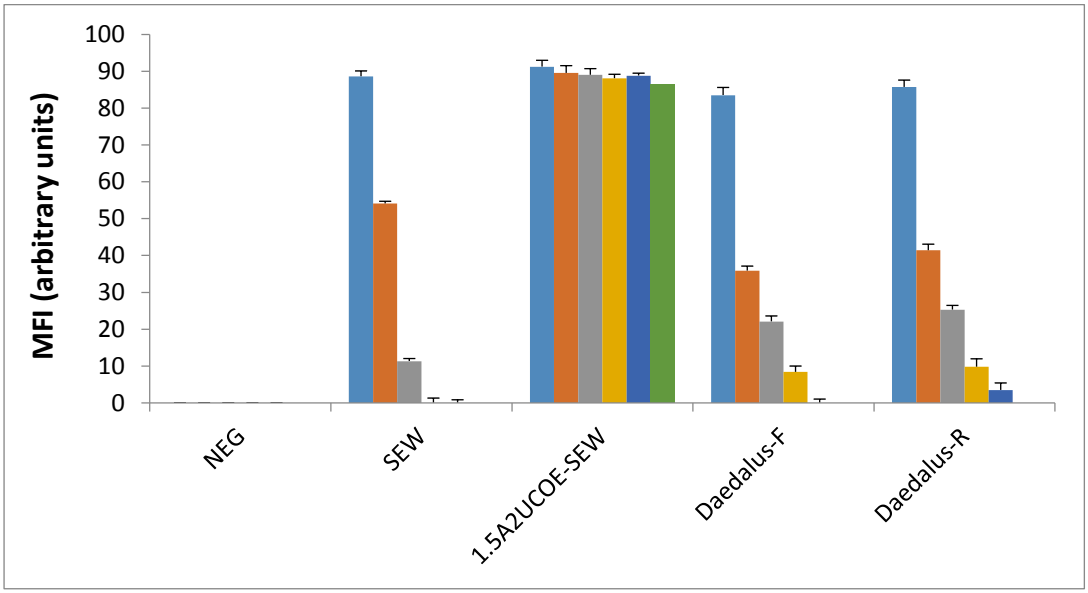


Figure 4.6. Novel candidate UCOEs Daedalus-F and Daedalus-R offer only partial protection against silencing in differentiated P19 cells. P19 cells were transduced with novel candidate UCOE Daedalus-F and Daedalus-R, control SEW and 1.5A2UCOE-SEW LVs (Figure 4.3). Cells were analysed by flow cytometry to detect percentage of eGFP reporter gene expressing (eGFP+) cells, mean fluorescence intensity (MFI) and by real-time Q-PCR for average vector copy number (VCN) per cell. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 38 days post-transduction. (**A**) timecourse of percentage eGFP-positive cells; (Mean + SEM, n=4; **p<0.01). (**B**) As in (A) but showing MFI; (Mean + SEM, n=4; **p<0.01). (**C**) As in (A)/(B) but average VCN/cell; (Mean + SEM, n=4; **p<0.01).

A)


EGFP-Positive Cells (%)
50
45
40
35
30
25
20
15
10
5
0
NEG
SEW
1.5A2UCOE-SEW
Daedalus-F
Daedalus-R
Day 3
DAY 10
DAY 17
DAY 24
DAY 31
Day 38

B)


MFI (arbitrary units)
100
90
80
70
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
Daedalus-F
Daedalus-R

C)

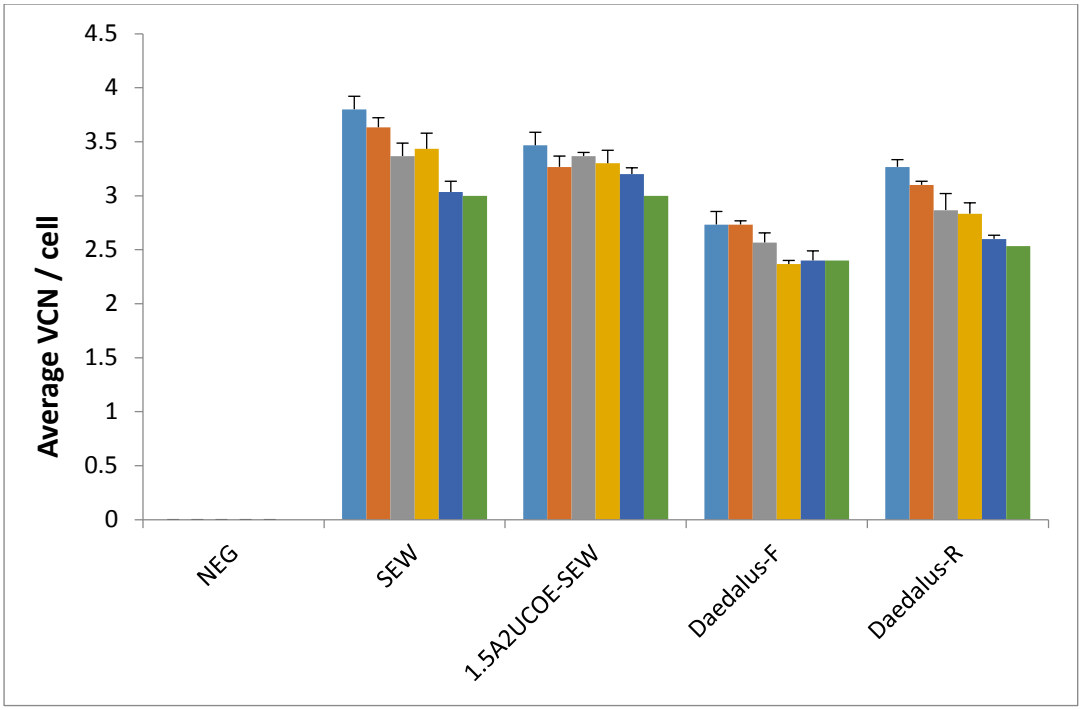


Figure 4.7. Novel candidate UCOEs Daedalus-F and Daedalus-R offer only partial protection against silencing in differentiated F9 cells. F9 cells were transduced with novel candidate UCOE Daedalus-F, Daedalus-R, control SEW and 1.5A2UCOE-SEW LVs (Figure 4.3). Cells were analysed by flow cytometry to detect percentage of eGFP reporter gene expressing (eGFP+) cells, mean fluorescence intensity (MFI) and by real-time Q-PCR for average vector copy number (VCN) per cell. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 38 days post-transduction. (**A**) timecourse of percentage eGFP-positive cells; (Mean + SEM, n=4; **p<0.01). (**B**) As in (A) but showing MFI; (Mean + SEM, n=4; **p<0.01). (**C**) As in (A)/(B) but average VCN/cell; (Mean + SEM, n=4; **p<0.01).

Because the both Daedalus constructs have not shown any stability, we decided not to make immunofluorescence staining counting of the differentiated cells.

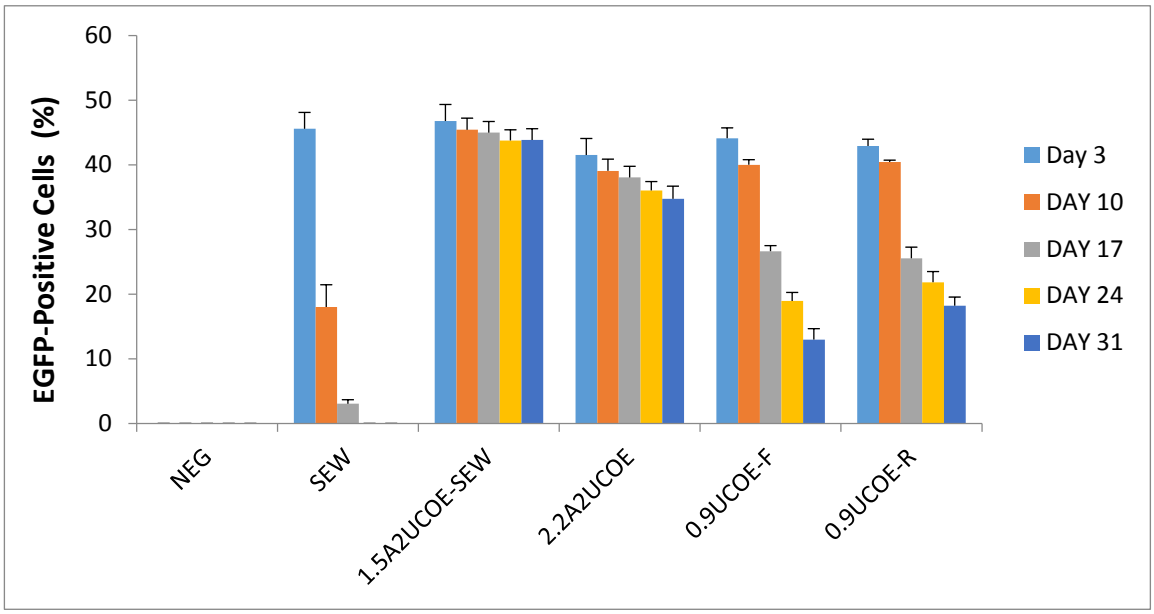
4.6. Functional analysis of candidate 0.9UCOE-F and 0.9UCOE-R in undifferentiated P19 and F9 cells

P19 and F9 cells were transduced with the candidate 0.9UCOE-F and 0.9UCOE-R LVs (Figure 4.3) harbouring a putative MDR at MOI of 3 or 6, with the intention to start the experiment within the range of 40-60% eGFP-positive cells in all pools. The transduced cells and negative control were then propagated in culture and assayed as a time course by flow cytometry, to determine the percentage of eGFP positive cells and the mean fluorescence intensity (MFI) for each repeat transduction for each vector. Cell cultures were analysed every 7 days from day 3 post-transduction and extending to 31 days (Figure 4.8 and 4.9). In addition, DNA was extracted from cells at each time point for RT-qPCR in order to determine the average vector copy number per cell.

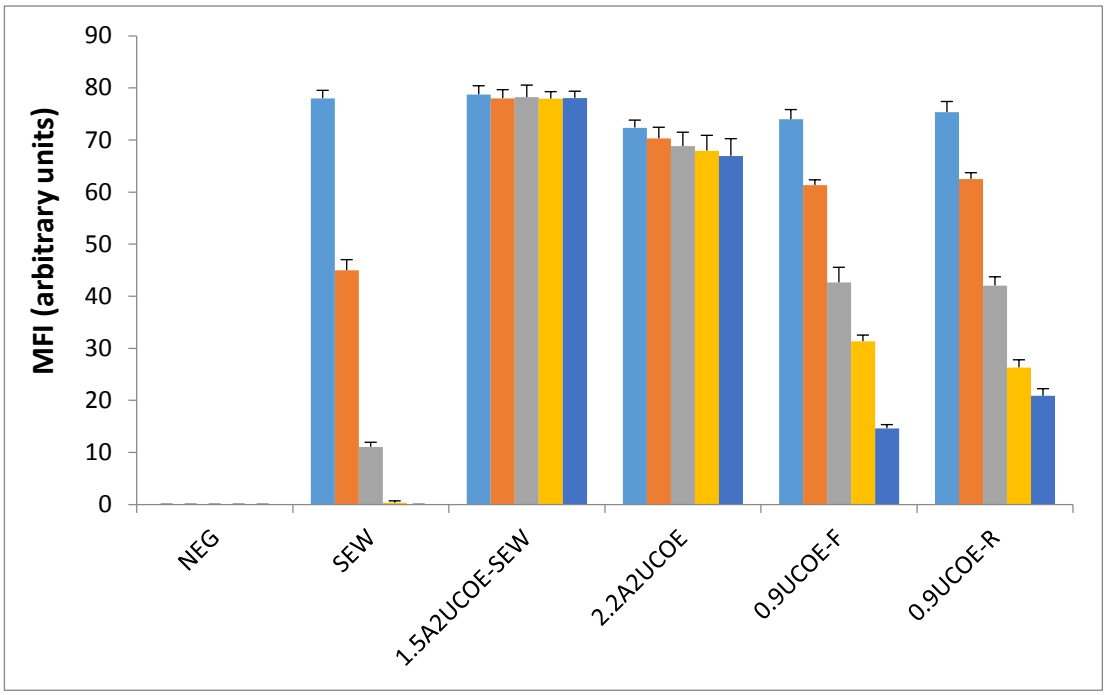
Figure 4.8A shows the flow cytometry time course results depicting percentage of eGFP+ cells. Although similar initial transduction efficiency was obtained with all vectors (45–60% eGFP+ cells), as we see reproducibly expression from the SEW (SFFV-eGFP-WPRE), rapidly declined from 45% to 2% positive cells within 17 days. In contrast, the proportion of eGFP+ cells from the 1.5A2UCOE-SEW and 2.2A2UCOE positive control vectors remained completely stable over the 31-day

period of culture. Expression from the 0.9UCOE-F and 0.9UCOE-R LVs where the 0.9kb core A2UCOE is linked to the SFFV promoter is also unstable but the rate of decrease in eGFP+ cells is markedly slower than that from SEW. The expression of eGFP in SEW transduced cells had dropped almost 80% after 2 weeks post-transduction. Contrastingly, the reduction we observed in eGFP+ cells for the 0.9UCOE-F and 0.9UCOE-R LVs was only 20% to 30% over the same time period. The values of mean fluorescent intensity (MFI) for all vectors paralleled the eGFP expression results (Figure 4.8B). MFI was stable in the 1.5A2UCOE-SEW transduced cells, unstable in the case of SEW and in the novel test 0.9UCOE-F and 0.9UCOE-R LV constructs. As consistently observed the average vector copy number per cell remained stable in all cases over the entire period of the experiment (Figure 4.8C), thus confirming that reduced levels of eGFP expression with the SEW, 0.9UCOE-F and 0.9UCOE-R LVs was due to silencing and not vector loss.

A)


EGFP-Positive Cells (%)
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
0.9UCOE-F
0.9UCOE-R
Day 3
DAY 10
DAY 17
DAY 24
DAY 31

B)


MFI (arbitrary units)
90
80
70
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
0.9UCOE-F
0.9UCOE-R

C)

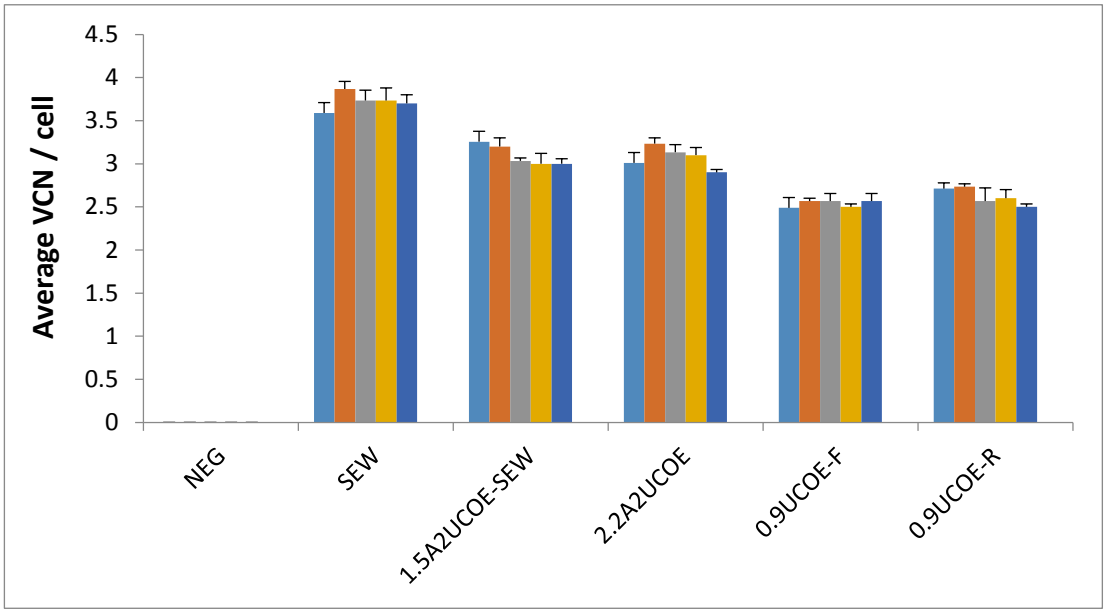
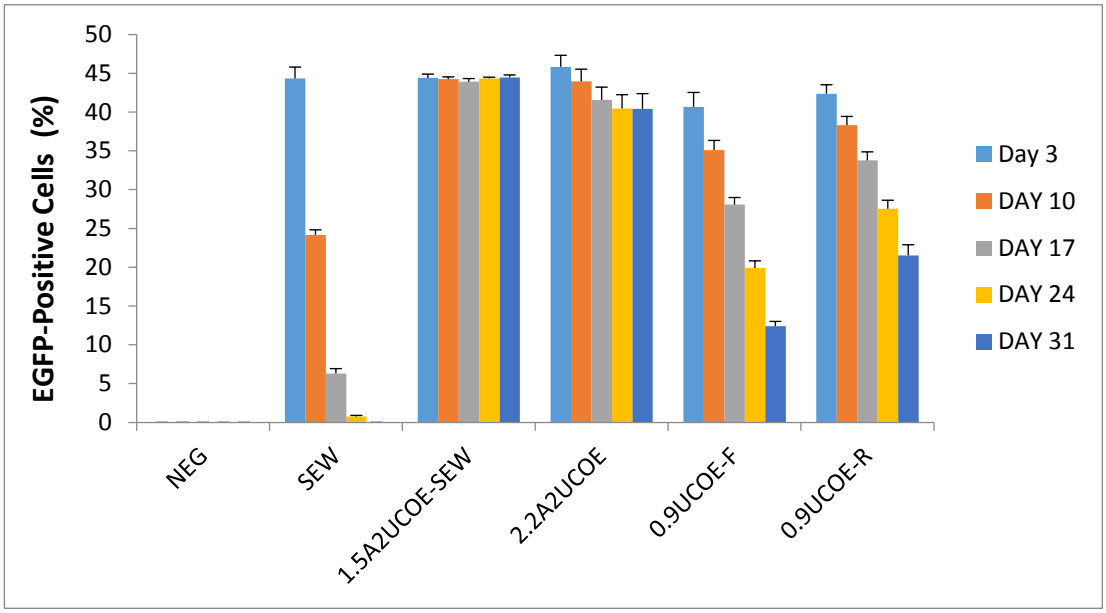


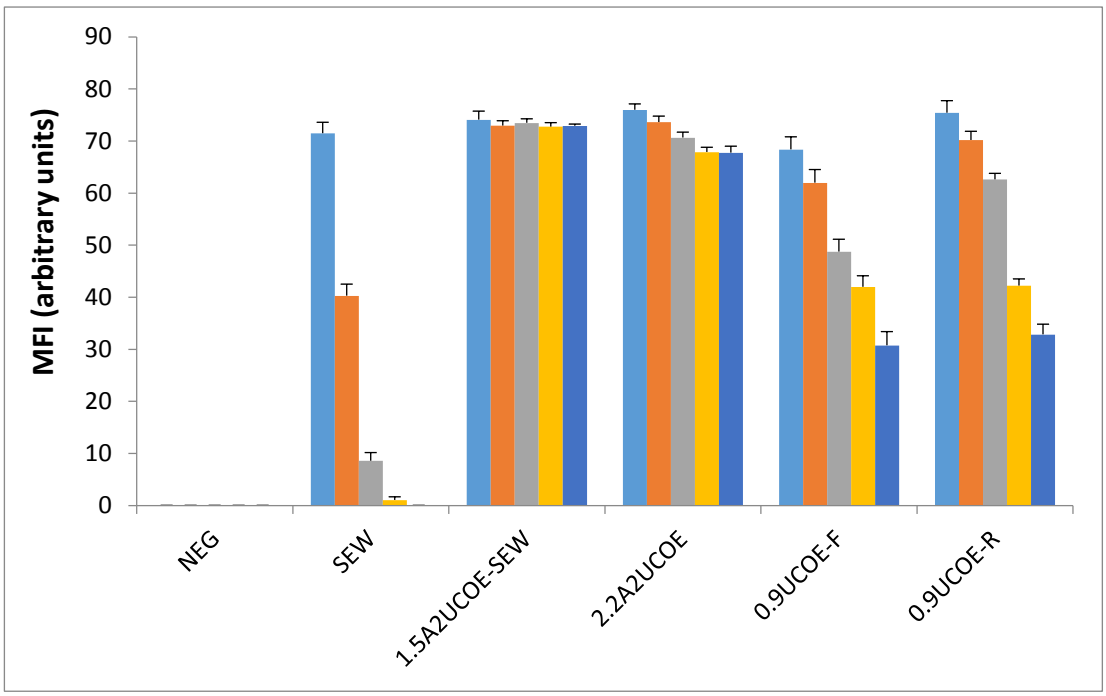
Figure 4.8. Novel candidate UCOEs 0.9UCOE-F and 0.9UCOE-R offer partial protection against silencing in undifferentiated P19 cells. P19 cells were transduced with the novel candidate UCOE constructs 0.9UCOE-F and 0.9UCOE-R and control SEW, 1.5A2UCOE-SEW and 2.2kbA2UCOE LVs (Figure 4.3). Cells were analysed by flow cytometry to detect percentage of eGFP reporter gene expressing (eGFP+) cells, mean fluorescence intensity (MFI) and by RT-Q-PCR for average vector copy number (VCN) per cell. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 31 days post-transduction. (**A**) timecourse of percentage eGFP-positive cells; (Mean + SEM, n=4; **p<0.01). (**B**) As in (A) but showing MFI; (Mean + SEM, n=4; **p<0.01). (**C**) As in (A)/(B) but average VCN/cell; (Mean + SEM, n=4; **p<0.01).

Figure 4.9A shows the flow cytometry time course results depicting percentage of eGFP+ cells in undifferentiated F9 cells transduced with the same 0.9UCOE-F and 0.9UCOE-R LVs. The results show a parallel activity to that we observed in P19 cells (Figure 4.8A). Although similar initial transduction efficiency was obtained with all vectors (39–50% eGFP+ cells), expression from SEW (SFFV-eGFP-WPRE) rapidly declined from 44% to 3% positive cells within 17 days. As in the case with P19 cells (Figure 4.8A), the proportion of eGFP+ cells from the 1.5A2UCOE-SEW and 2.2A2UCOE vectors remained completely stable over the 31-day period of culture. Again, expression from the candidate 0.9UCOE-F and 0.9UCOE-R LVs also declined over time but at significantly slower rates compared with SEW. The expression of eGFP in SEW transduced cells had dropped almost 80% after 2 weeks post-transduction. Contrastingly, the reduction observed in eGFP+ cells was only 5-10% for the 0.9UCOE-F and 0.9UCOE-R LVs over the same time period. The values of mean fluorescent intensity (MFI) for all vectors paralleled the eGFP expression results (Figure 4.9B). MFI was stable in the 1.5A2UCOE-SEW and 2.2A2UCOE transduced cells, unstable in the case of SEW and partially stable in the novel test UCOE constructs. Analysis of average vector copy number per cell revealed that this was stable in all cases over the entire time period of the experiment (Figure 4.9C).

A)


EGFP-Positive Cells (%)
50
45
40
35
30
25
20
15
10
5
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
0.9UCOE-F
0.9UCOE-R
Day 3
DAY 10
DAY 17
DAY 24
DAY 31

B)


MFI (arbitrary units)
90
80
70
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
0.9UCOE-F
0.9UCOE-R

C)

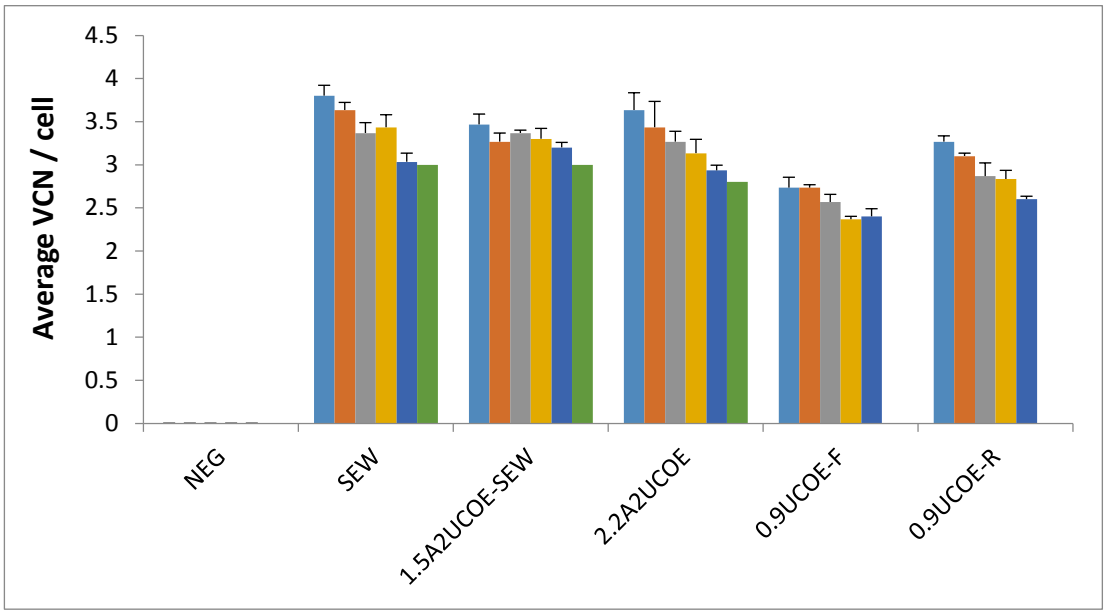


Figure 4.9. Novel candidate UCOEs 0.9UCOE-F and 0.9UCOE-R offer partial protection against silencing in undifferentiated F9 cells. F9 cells were transduced with novel candidate UCOE 0.9UCOE-F and 0.9UCOE-R, and control SEW, 1.5A2UCOE-SEW and 2.2A2UCOE LVs (Figure 4.3). Cells were analysed by flow cytometry to detect percentage of eGFP reporter gene expressing (eGFP+) cells, mean fluorescence intensity (MFI) and by RT-Q-PCR for average vector copy number (VCN) per cell. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 31 days post-transduction. (**A**) timecourse of percentage eGFP-positive cells; (Mean + SEM, n=4; **p<0.01). (**B**) As in (A) but showing MFI; (Mean + SEM, n=4; **p<0.01). (**C**) As in (A)/(B) but average VCN/cell; (Mean + SEM, n=4; **p<0.01).

4.7 Functional assay of candidate 0.9UCOE-F and 0.9UCOE-R UCOEs in differentiated P19 and F9 cells

The novel candidate UCOE 0.9UCOE-F and 0.9UCOE-R vectors were able to confer partial stability of expression in undifferentiated P19 (Figure 4.8) and F9 (Figure 4.9) cells. We next evaluated the ability of these novel LVs to stabilize expression upon differentiation of P19 and F9 cells down the neuroectodermal and endodermal lineages, respectively.

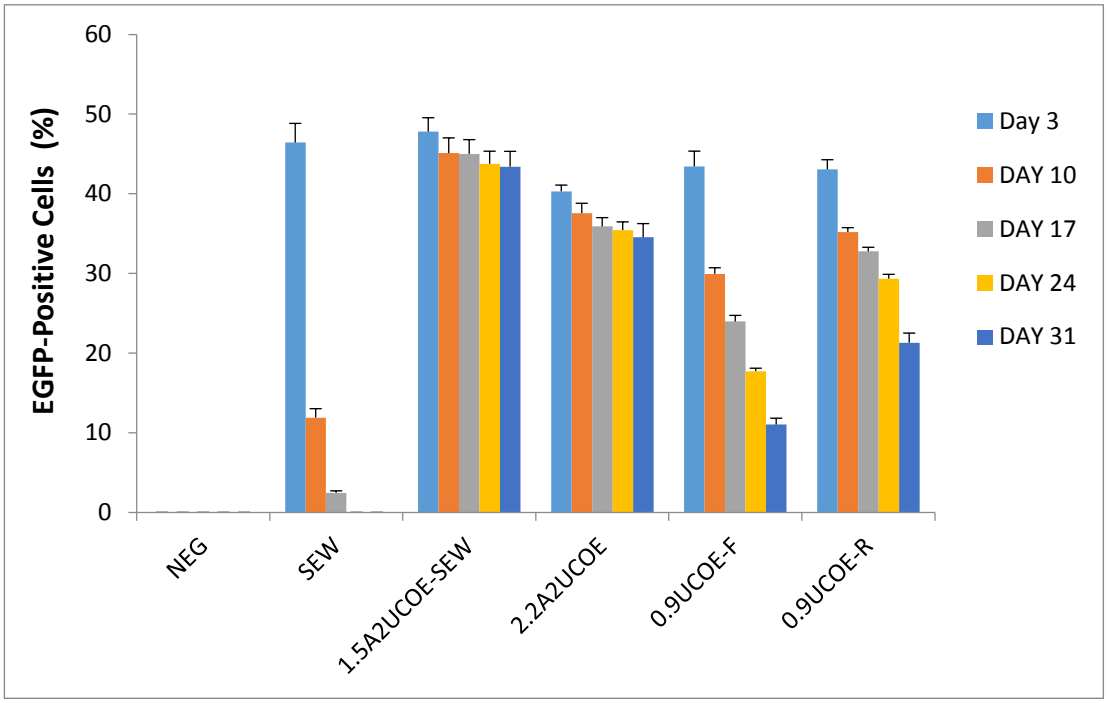
P19 and F9 cells transduced with each LV were first induced to form embryoid bodies in retinoic acid-containing differentiation medium and subsequent neuroectodermal and parietal endoderm cell differentiation, respectively as before.

The expression results from triplicate samples of cells transduced with the SEW, 1.5A2UCOE-SEW, 2.2A2UCOE, 0.9UCOE-F and 0.9UCOE-R LVs following differentiation is presented in Figures 4.10 and 4.11. Our data show that the control LVs carrying the positive UCOE constructs (1.5A2UCOE-SEW and 2.2A2UCOE) retained stable expression after P19 (Figure 4.10A) and F9 (Figure 4.11A) cell differentiation. However, the novel candidate UCOE vectors 0.9UCOE-F and 0.9UCOE-R show a gradual reduction in eGFP expression whilst eGFP in the silencing control SEW vector was rapidly repressed (Figures 4.10A and 4.11A).

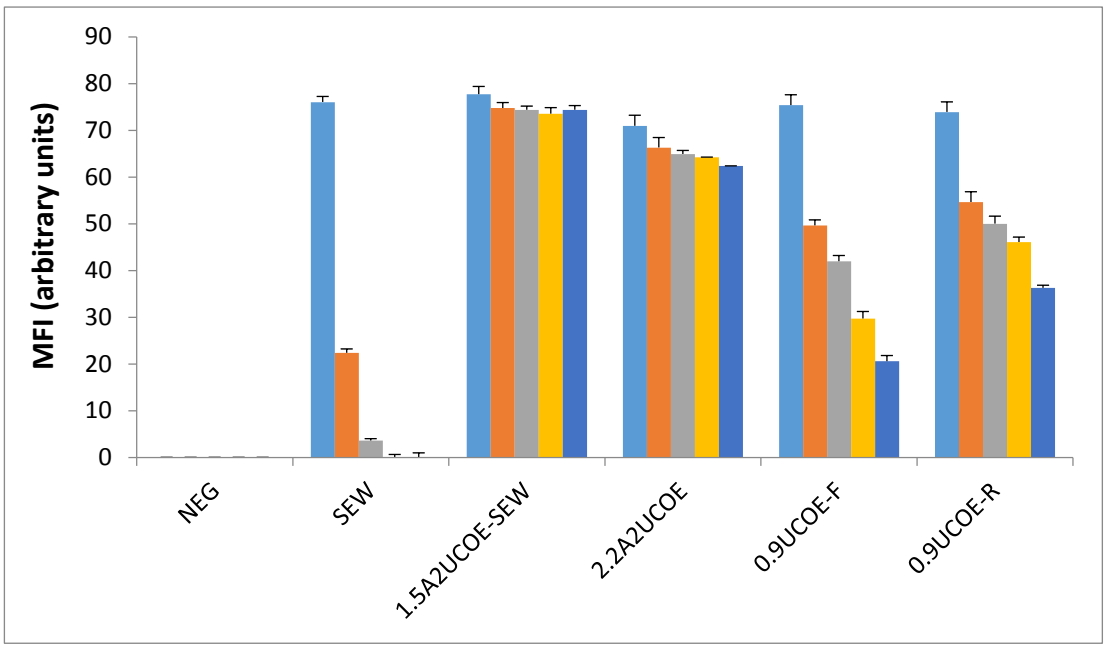
As in the case of undifferentiated cells (Figures 4.8 and 4.9), MFI values upon differentiation of both P19 and F9 cells paralleled the eGFP+ cell expression data (Figures 4.10B and 4.11B) whilst average vector copy number was essentially stable throughout the timecourse of the experiment (Figures 4.10C and 4.11C).

Thus the LVs tested in these experiments gave similar outcomes in undifferentiated and differentiated F9 and P19 cells. Interestingly, the rate of reduction in eGFP expression with the SEW and novel candidate 0.9UCOE-F and 0.9UCOE-R LVs is greater upon differentiation (Figures 4.10A and 4.11A) than in undifferentiated cells (Figures 4.8A and 4.9A) cells.

A)


EGFP-Positive Cells (%)
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
0.9UCOE-F
0.9UCOE-R
Day 3
DAY 10
DAY 17
DAY 24
DAY 31

B)


MFI (arbitrary units)
90
80
70
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
0.9UCOE-F
0.9UCOE-R

C)

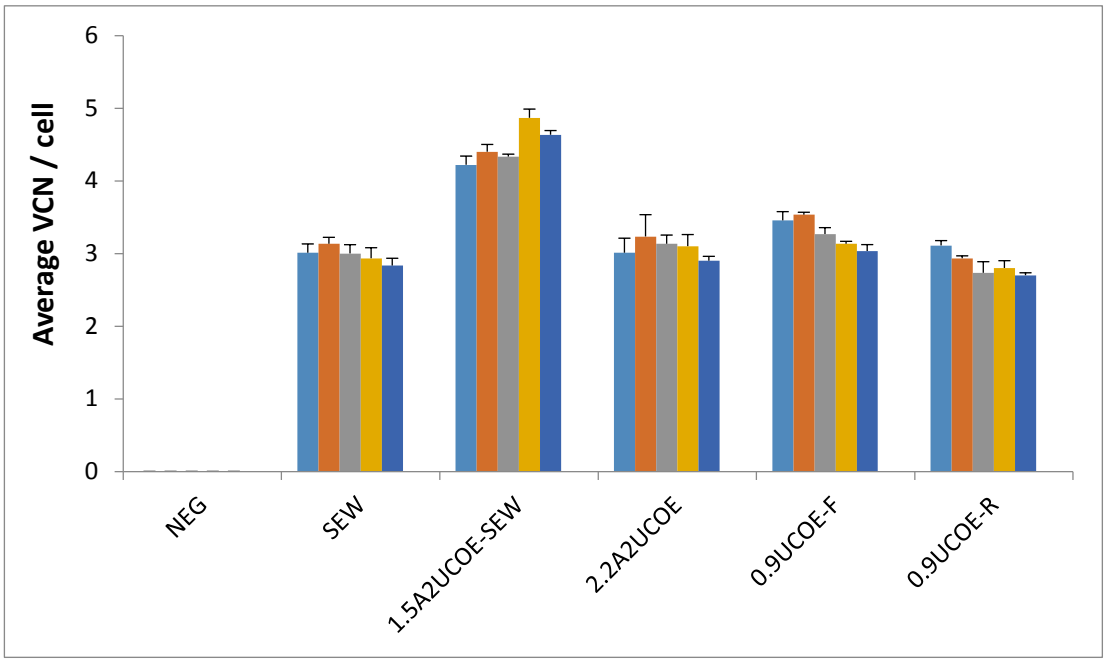
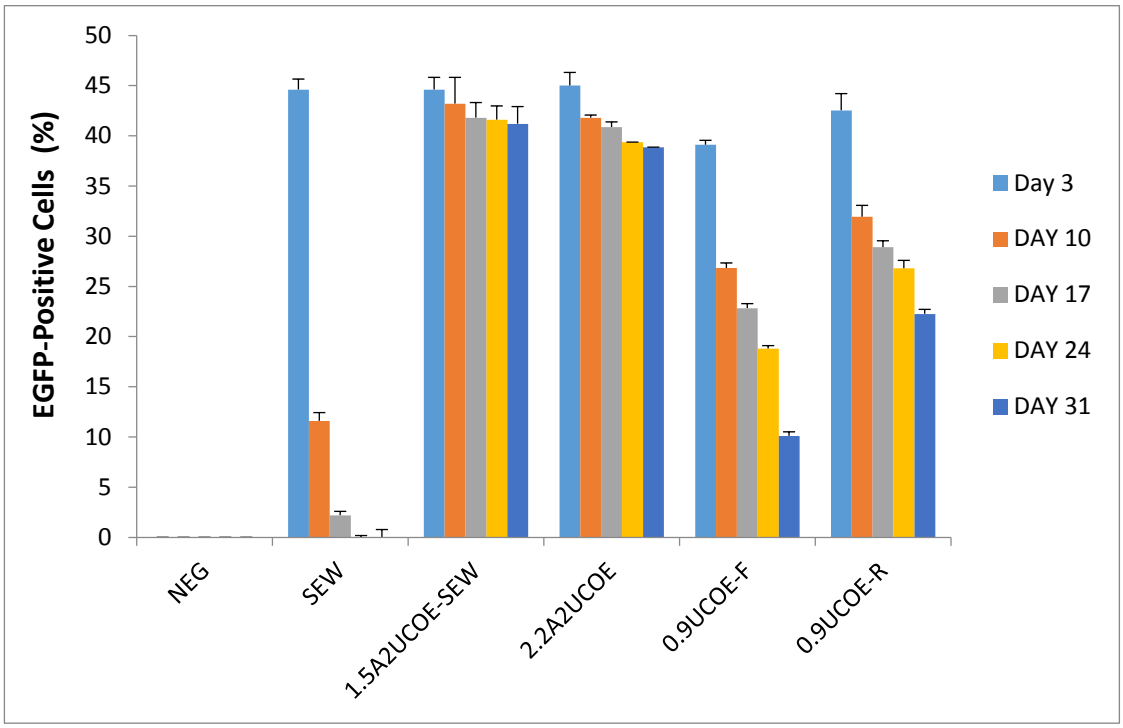
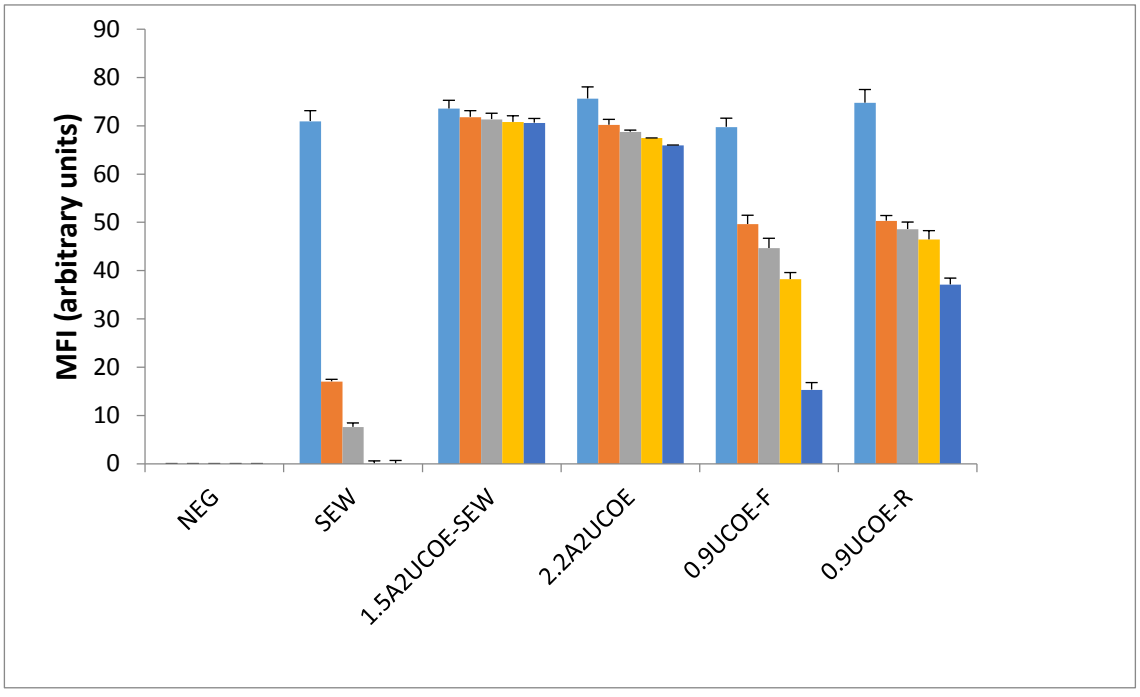


Figure 4.10. Novel candidate 0.9UCOE-F and 0.9UCOE-R UCOEs offer partial protection against silencing in differentiated P19 cells. P19 cells were transduced with novel candidate UCOE 0.9UCOE-F and 0.9UCOE-R, and control SEW, 1.5A2UCOE-SEW and 2.2A2UCOE LVs (Figure 4.3). Cells were analysed by flow cytometry to detect percentage of eGFP reporter gene expressing (eGFP+) cells, mean fluorescence intensity (MFI) and by RT-Q-PCR for average vector copy number (VCN) per cell. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 31 days post-transduction. (**A**) timecourse of percentage eGFP-positive cells; (Mean + SEM, n=4; **p<0.01). (**B**) As in (A) but showing MFI; (Mean + SEM, n=4; **p<0.01). (**C**) As in (A)/(B) but average VCN/cell; (Mean + SEM, n=4; **p<0.01).

A)


EGFP-Positive Cells (%)
50
45
40
35
30
25
20
15
10
5
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
0.9UCOE-F
0.9UCOE-R
Day 3
DAY 10
DAY 17
DAY 24
DAY 31

B)


MFI (arbitrary units)
90
80
70
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
0.9UCOE-F
0.9UCOE-R

C)

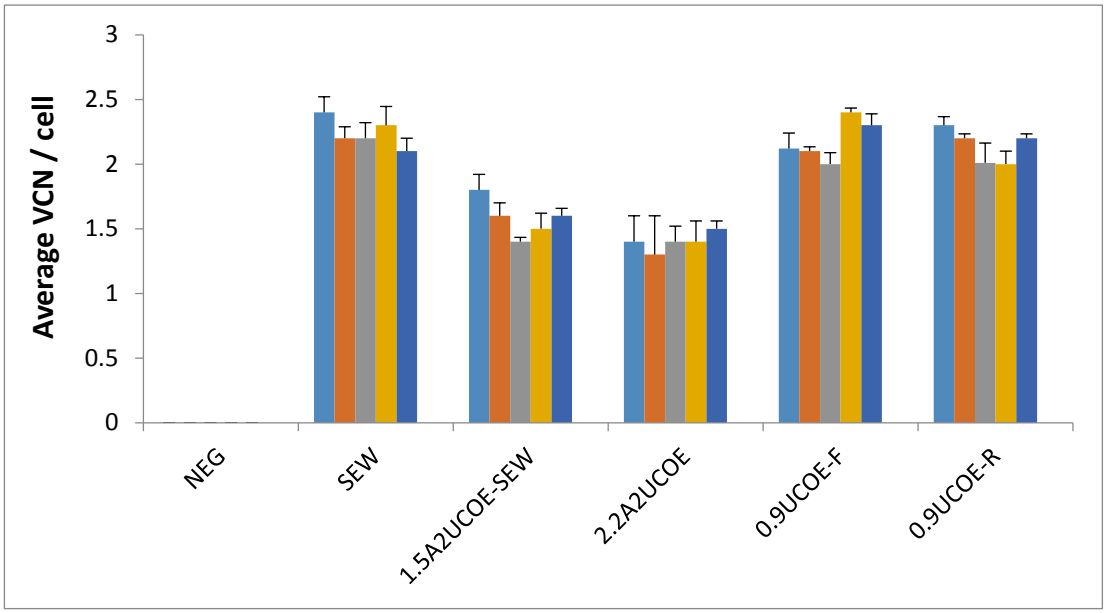


Figure 4.11. Novel candidate 0.9UCOE-F and 0.9UCOE-R UCOEs offer partial protection against silencing in differentiated F9 cells. F9 cells were transduced with novel candidate UCOE 0.9UCOE-F and 0.9UCOE-R, and control SEW, 1.5A2UCOE-SEW and 2.2A2UCOE LVs (Figure 4.3). Cells were analysed by flow cytometry to detect percentage of eGFP reporter gene expressing (eGFP+) cells, mean fluorescence intensity (MFI) and by RT-Q-PCR for average vector copy number (VCN) per cell. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 31 days post-transduction. (**A**) timecourse of percentage eGFP-positive cells; (Mean + SEM, n=4; **p<0.01). (**B**) As in (A) but showing MFI; (Mean + SEM, n=4; **p<0.01). (**C**) As in (A)/(B) but average VCN/cell; (Mean + SEM, n=4; **p<0.01).

4.8 Evaluation of novel candidate 0.9UCOE-F and 0.9UCOE-R UCOE LVs in differentiated P19 and F9 cells by immunofluorescence staining

In order to confirm the results obtained by flow cytometry with differentiated P19 and F9 cells transduced with the novel candidate 0.9UCOE-F and 0.9UCOE-R UCOE LVs (Figures 4.10 and 4.11), P19 and F9 cells which had undergone neuroectodermal and endodermal differentiation, respectively, were stained for appropriate markers and scored under immunofluorescence microscopy (see section 3.7). The results depicting transduced GFP plus β-tubulin III double-positive differentiated P19 cells are shown in Figure 4.12 and GFP plus Oct3-4 double-positive differentiated F9 cells in Figure 4.13. These data corroborate the results obtained by flow cytometry (Figures 4.10 and 4.11) and show that the 0.9UCOE-F and 0.9UCOE-R candidate UCOEs confer only marginal protection against silencing of the linked SFFV promoter over during the course of differentiation.

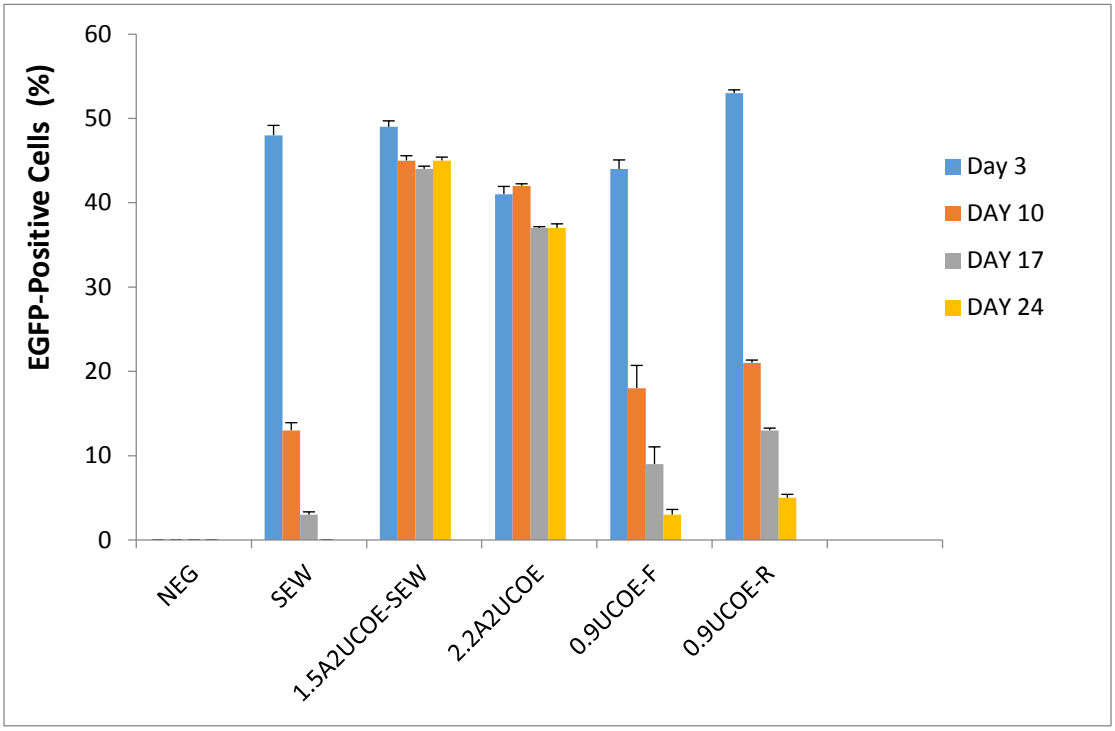


Figure 4.12: Novel candidate UCOEs offer only partial protection against silencing in differentiated P19 cells. P19 cells were transduced with novel candidate UCOE 0.9UCOE-F and 0.9UCOE-R, and control SEW, 1.5A2UCOE-SEW and 2.2A2UCOE LVs (Figure 4.3) and at 3-days post-transduction induced to undergo differentiation down the neuroectodermal lineage. Cells were analysed at various times following differentiation by fluorescence microscopy by scoring for GFP and anti-β-tubulin III double positive cells. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 24 days post-transduction (Mean + SEM, n=4; **p<0.01).

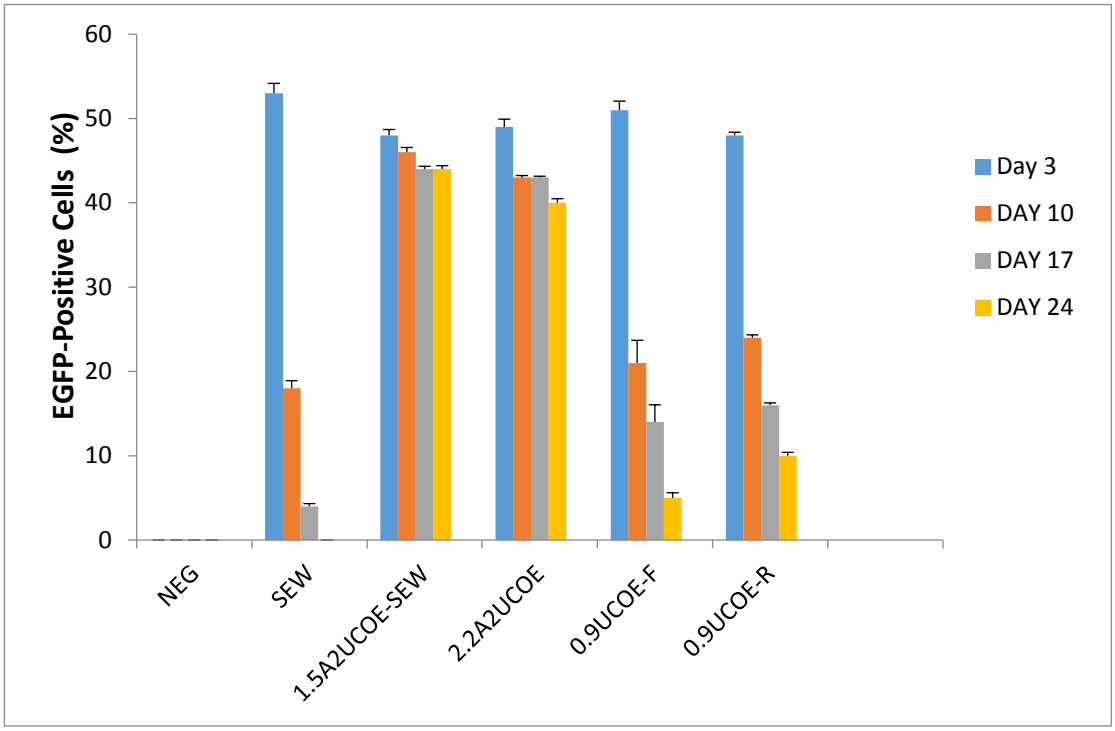


Figure 4.13: Novel candidate UCOEs offer only partial protection against silencing in differentiated F9 cells. F9 cells were transduced with novel candidate 0.9UCOE-F and 0.9UCOE-R and control SEW, 1.5A2UCOE-SEW and 2.2A2UCOE LVs (Figure 4.3) and at 3-days post-transduction induced to undergo differentiation down the endodermal lineage. Cells were analysed at various times following differentiation by fluorescence microscope by scoring for GFP and anti-Oct3-4 double positive cells. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 24 days post-transduction (Mean + SEM, n=4; **p<0.01).

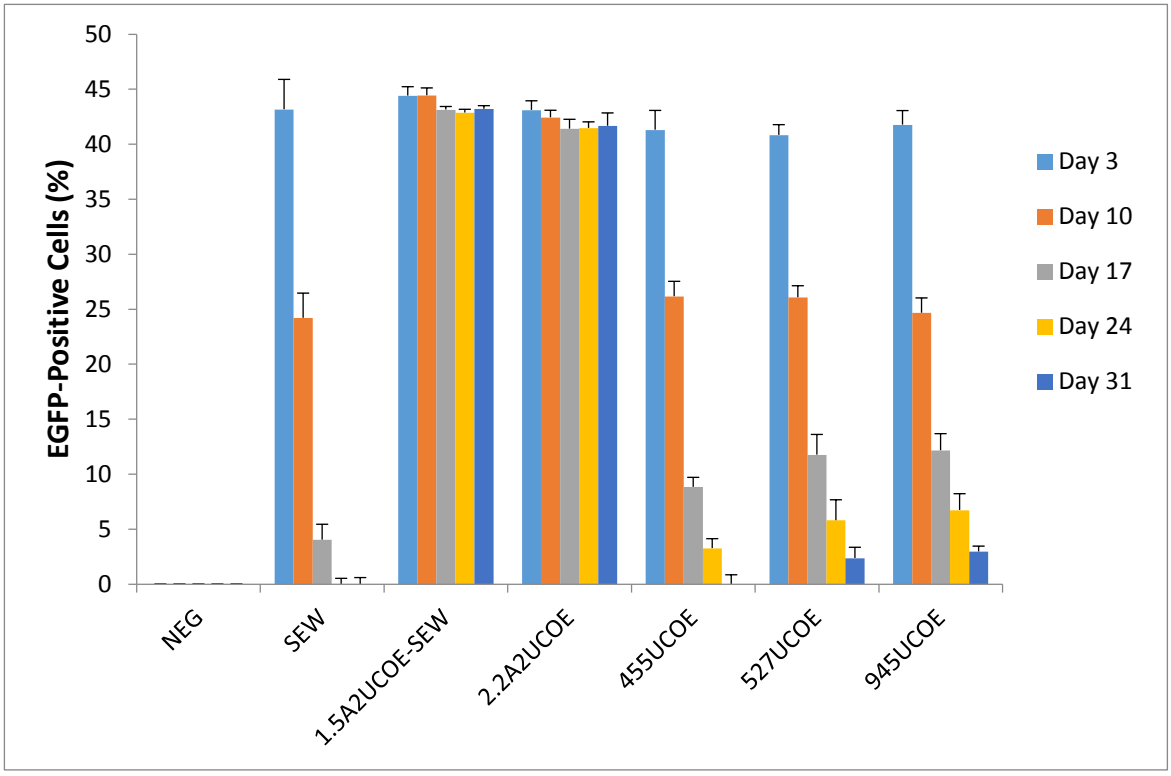
4.9 Functional assay of candidate 455UCOE, 527UCOE and 945UCOE LVs in undifferentiated P19 and F9 cells

P19 and F9 cells were transduced with the novel candidate 455UCOE, 527UCOE and 945UCOE LVs at a MOI of 3 or 6, with the intention to start the experiment within the range of 40-60% eGFP+ cells in all pools. The transduced cells and negative control were then propagated in culture and assayed as a time course by flow cytometry, to determine the percentage of eGFP-positive cells and the mean fluorescence intensity (MFI) for each repeat transduction for each vector. Cell cultures were analysed every 7 days from day 3 post-transduction and extending to 31 days (Figure 4.12 and 4.13). In addition, DNA was extracted from cells at each time point for RT-qPCR in order to determine the average vector copy number per cell.

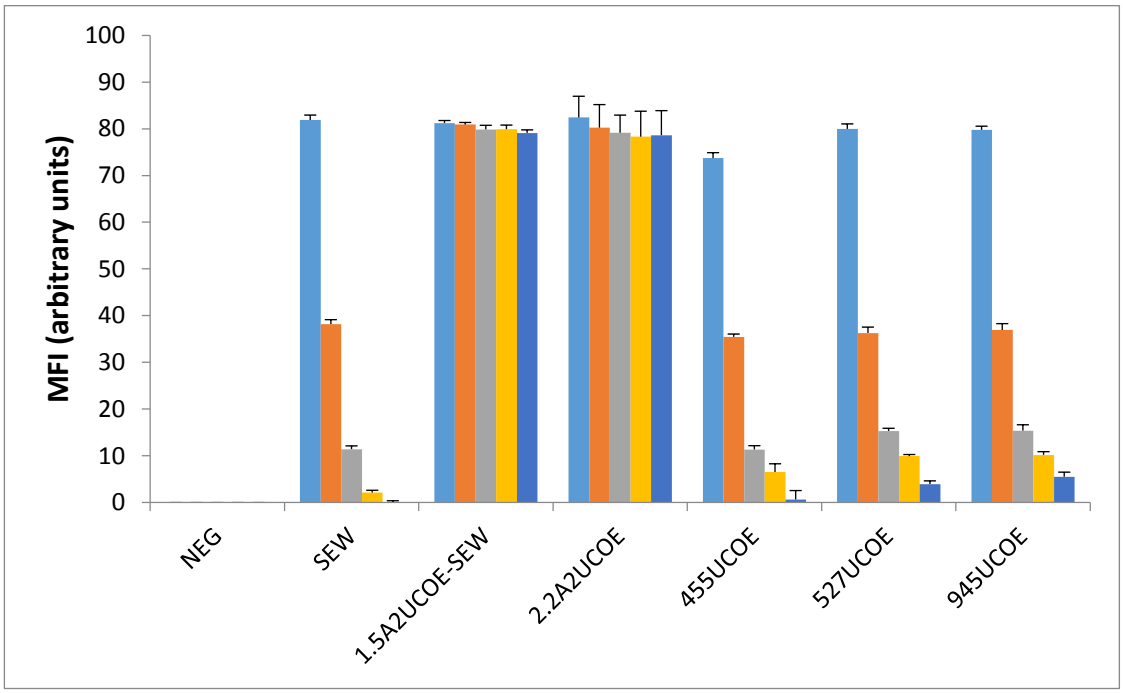
Figure 4.12A shows the flow cytometry time course results depicting percentage of eGFP+ cells. A similar initial transduction efficiency was obtained with all vectors (45–60% eGFP+ cells). As expected expression from SEW (SFFV-eGFP-WPRE), rapidly declined from 43% to 2% eGFP+ positive cells within 17 days whilst the proportion of eGFP-positive cells from the 1.5A2UCOE-SEW and 2.2A2UCOE vectors remained completely stable over the 31-day period of culture. The novel candidate 455UCOE, 527UCOE and 945UCOE LVs also resulted in a rapid decrease of eGFP expression in parallel to that seen with SEW. The expression of eGFP with SEW transduced cells dropped almost 80% after 2 weeks post-transduction. Similarly, a 60-70% reduction in eGFP-positive cells occurred with the test 455UCOE, 527UCOE and 945UCOE LVs over the same time period. The MFI values for all vectors matched the eGFP expression results (Figure 4.12B). MFI was stable in the

1.5A2UCOE-SEW and 2.2A2UCOE transduced cells and unstable for both the SEW and 455UCOE, 527UCOE and 945UCOE test constructs.

A)


EGFP-Positive Cells (%)
50
45
40
35
30
25
20
15
10
5
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
455UCOE
527UCOE
945UCOE
Day 3
Day 10
Day 17
Day 24
Day 31

B)


MFI (arbitrary units)
100
90
80
70
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
455UCOE
527UCOE
945UCOE

C)

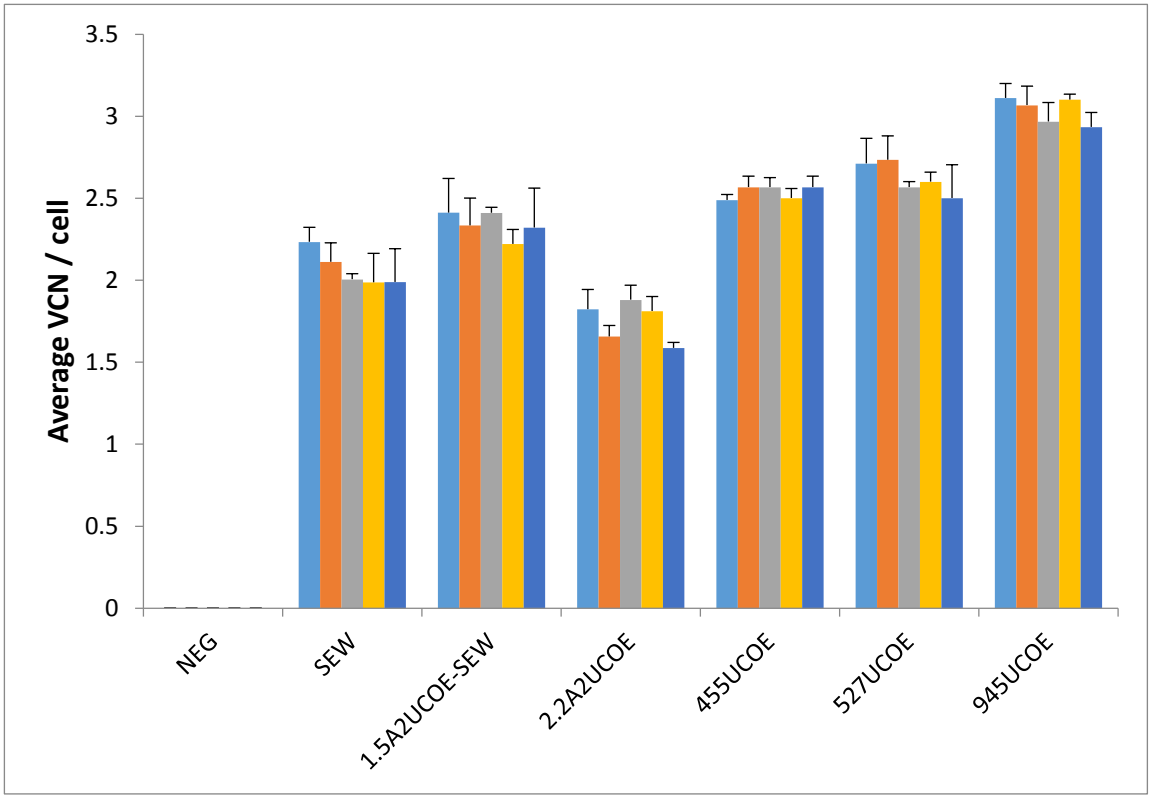
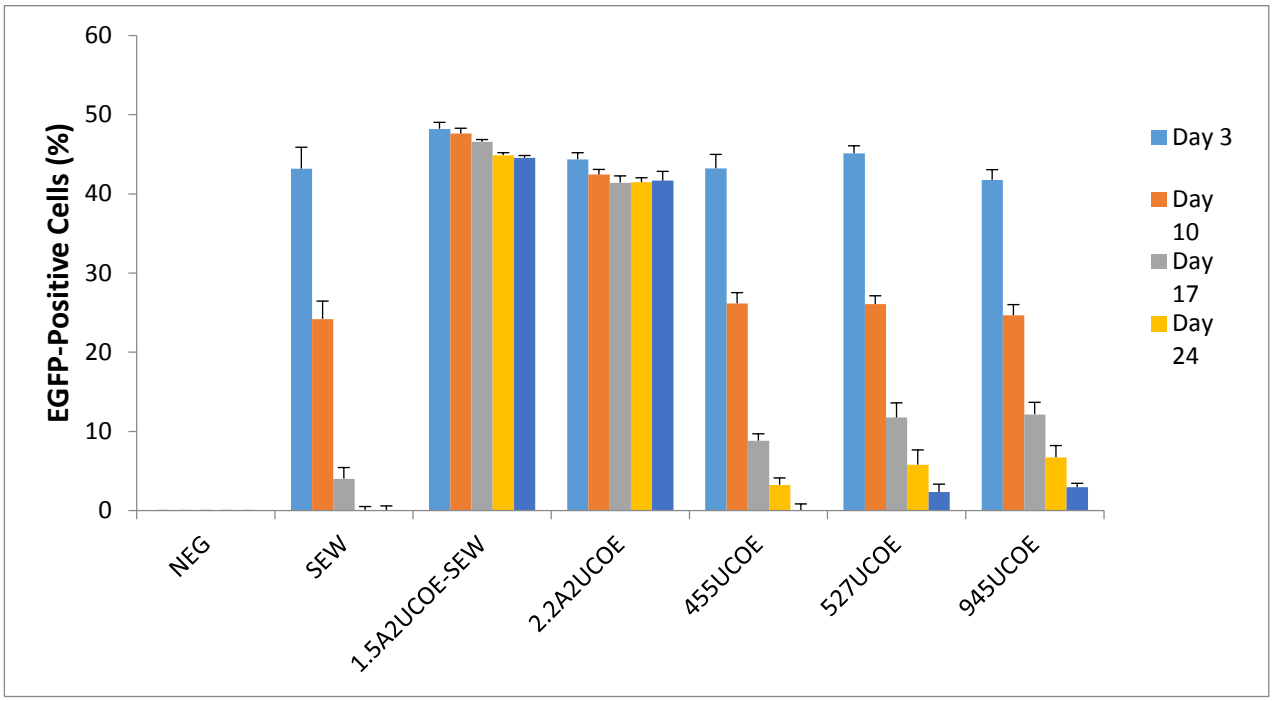


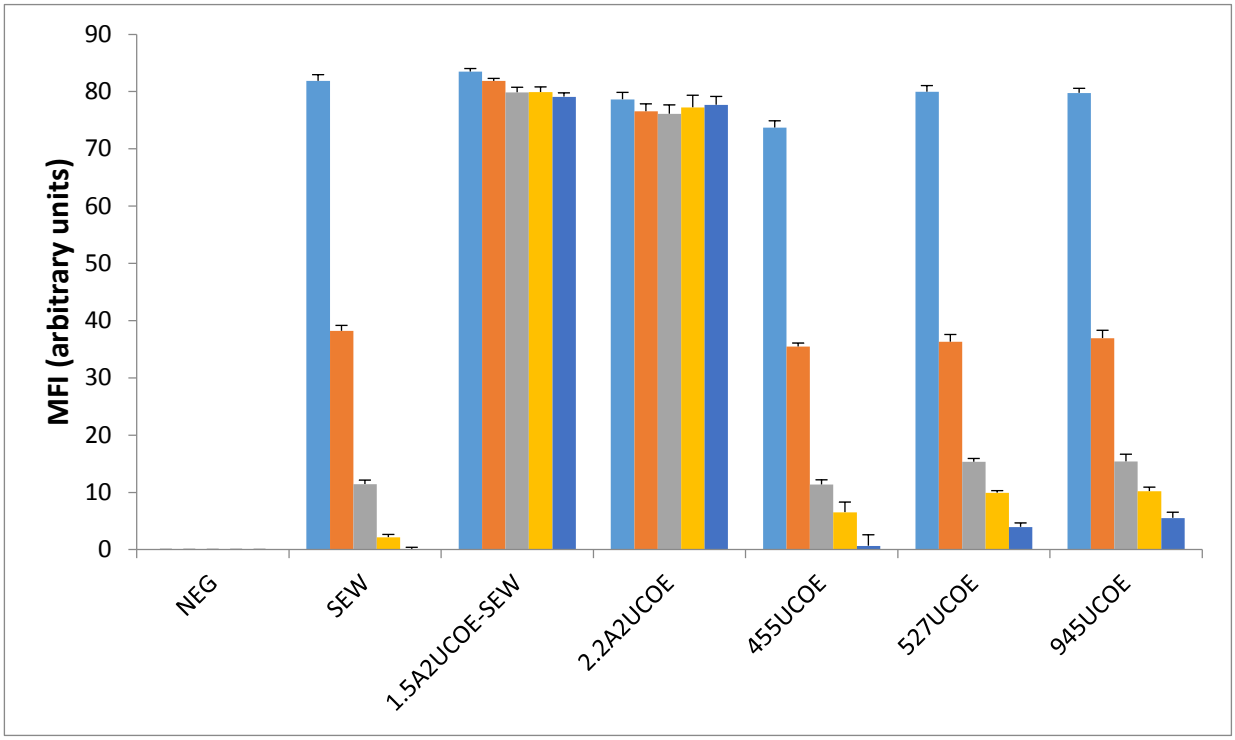
Figure 4.14. Novel candidate 455UCOE, 527UCOE and 945UCOE do not confer protection against silencing in undifferentiated P19 cells. P19 cells were transduced with novel candidate 455UCOE, 527UCOE and 945UCOE constructs as well as control SEW, 1.5A2UCOE-SEW and 2.2A2UCOE LVs (Figure 4.3). Cells were analysed by flow cytometry to detect percentage of eGFP reporter gene expressing (eGFP+) cells, mean fluorescence intensity (MFI) and by RT-Q-PCR for average vector copy number (VCN) per cell. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 31 days post-transduction. (**A**) timecourse of percentage eGFP-positive cells; (Mean + SEM, n=4; **p<0.01). (**B**) As in (A) but showing MFI; (Mean + SEM, n=4; **p<0.01). (**C**) As in (A)/(B) but average VCN/cell; (Mean + SEM, n=4; **p<0.01).

Figure 4.13A shows the flow cytometry time course results from undifferentiated F9 cells transduced with the 455UCOE, 527UCOE and 945UCOE LVs. The results are very similar to what we observed in undifferentiated P19 cells (Figure 4.12A). Following transduction to give a similar initial level of eGFP+ cells in all cases (39–50% eGFP+ cells), expression of eGFP from SEW (SFFV-eGFP-WPRE) gave the usual rapid decline from 47% to 2% positive cells within 17 days (Figure 4.13A). Again, as expected, the proportion of eGFP+ cells from the positive control 1.5A2UCOE-SEW and 2.2A2UCOE vectors remained completely stable over the 38-day period of culture. Expression from the candidate 455UCOE, 527UCOE and 945UCOE vectors as in the case of P19 cells (Figure 4.12A) again resulted in a comparable rapid decline in eGFP+ cells to that seen with SEW (Figure 4.13A). As we have consistently seen, the expression of eGFP in SEW transduced cells had dropped almost 80% after 2 weeks post-transduction. Similarly, we observed a 40-50% reduction in eGFP-positive cells with the candidate 455UCOE, 527UCOE and 945UCOE LVs over the same time period. The MFI values for all vectors paralleled the eGFP expression results (Figure 4.13B). MFI was stable in the 1.5A2UCOE-SEW transduced cells and unstable in the case of SEW and the novel 455UCOE, 527UCOE and 945UCOE test constructs. The average vector copy number per cell was essentially unchanged in all cases during the course of the experiment confirming that there had been no vector loss.

A)


EGFP-Positive Cells (%)
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
455UCOE
527UCOE
945UCOE
Day 3
Day 10
Day 17
Day 24

B)


MFI (arbitrary units)
90
80
70
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
455UCOE
527UCOE
945UCOE

C)

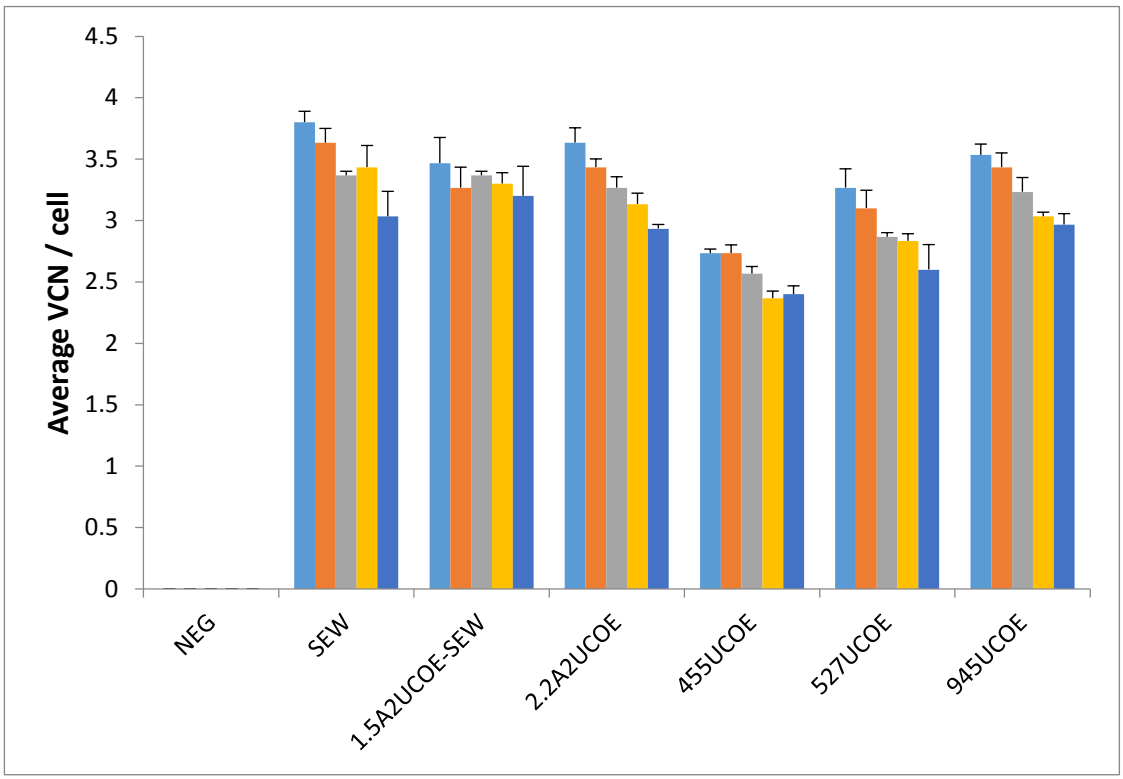


Figure 4.15. Novel candidate 455UCOE, 527UCOE and 945UCOE LVs do not confer protection against silencing in undifferentiated F9 cells. F9 cells were transduced with novel candidate 455UCOE, 527UCOE and 945UCOE constructs, and control SEW, 1.5A2UCOE-SEW and 2.2A2UCOE LVs (Figure 4.3). Cells were analysed by flow cytometry to detect percentage of eGFP reporter gene expressing (eGFP+) cells, mean fluorescence intensity (MFI) and by RT-Q-PCR for average vector copy number (VCN) per cell. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 31 days post-transduction. (**A**) timecourse of percentage eGFP-positive cells; (Mean + SEM, n=4; **p<0.01). (**B**) As in (A) but showing MFI; (Mean + SEM, n=4; **p<0.01). (**C**) As in (A)/(B) but average VCN/cell; (Mean + SEM, n=4; **p<0.01).

4.10 Functional analysis of candidate 455UCOE, 527UCOE and 945UCOE LVs in differentiated P19 and F9 cells

The novel candidate 455UCOE, 527UCOE and 945UCOE vectors were unable to confer stability of expression in undifferentiated P19 (Figure 4.12) and F9 (Figure 4.13) cells. In order to confirm and extend the data implying a lack of UCOE function from these constructs, we next evaluated their efficiency of expression during differentiation of P19 and F9 cells down the neuroectodermal and endodermal lineages, respectively.

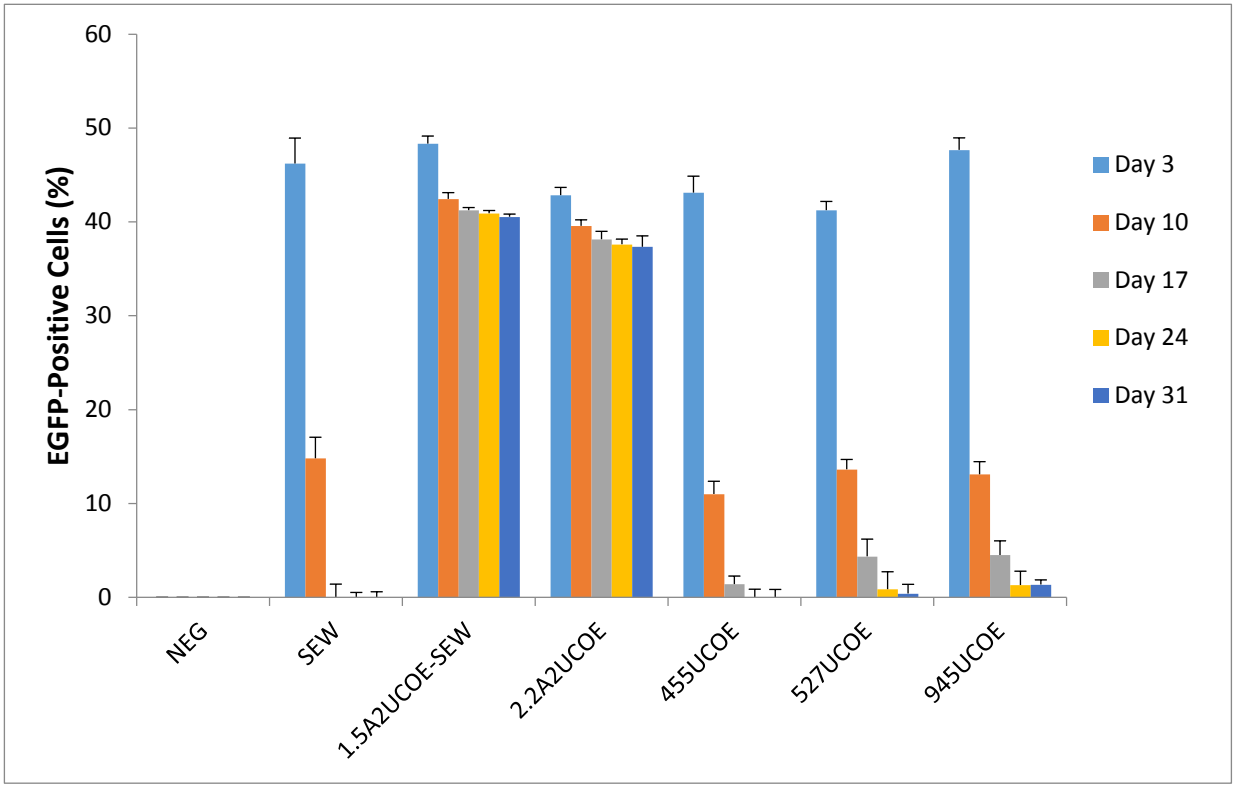
Our standard conditions were used to induce cells transduced with each LV to undergo differentiation via an embryoid body stage in the presence of retinoic acid-containing medium. The expression results from triplicate samples of cells transduced with the SEW, 1.5A2UCOE-SEW, 2.2A2UCOE, 455UCOE, 527UCOE and 945UCOE LVs following differentiation is presented in Figures 4.14 and 4.15. Our data show that the positive control LVs (1.5A2UCOE-SEW and 2.2A2UCOE) have retained stable expression after the P19 (Figure 4.14A) and F9 (Figure 4.15A) cells had undergone differentiation down the neuronal and endoderm lineages, respectively. However, the novel candidate 455UCOE, 527UCOE and 945UCOE vectors showed a similar rapid reduction in eGFP expression to that of the silencing control SEW LV (Figures 4.14A and 4.15A).

As in the case of undifferentiated cells (Figures 4.12B and 4.13B), MFI values upon differentiation of both P19 and F9 cells paralleled the eGFP+ cell expression data

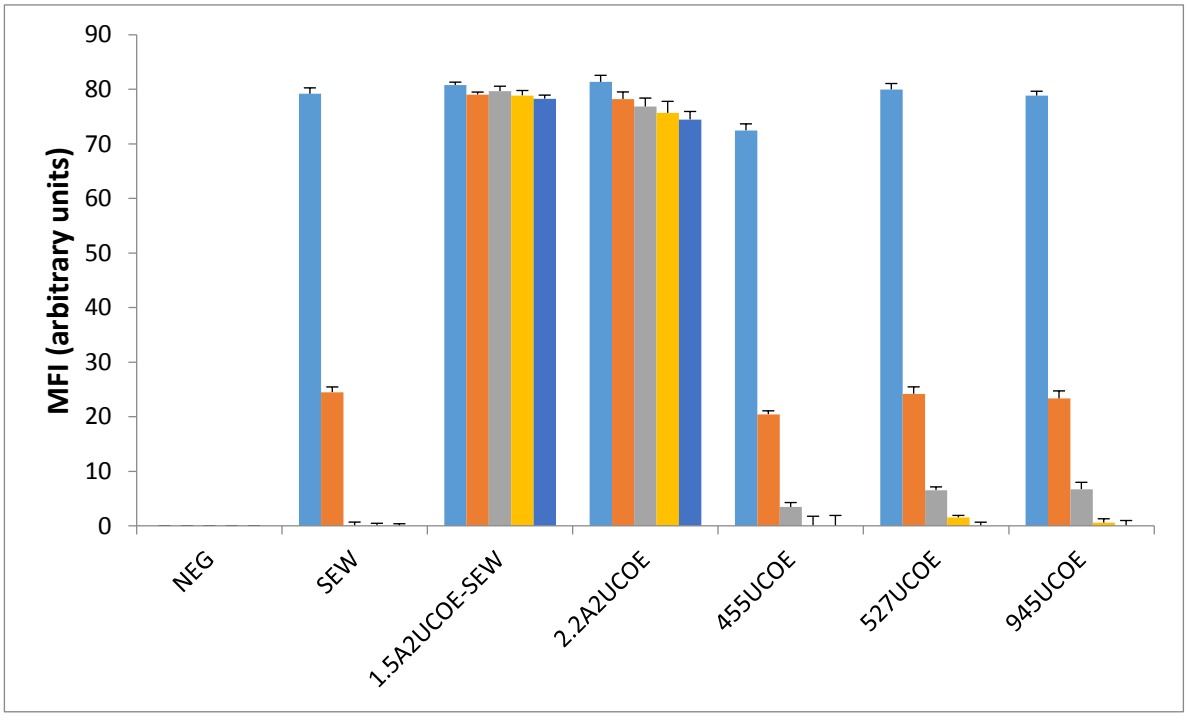
(Figures 4.14B and 4.15B) whilst average vector copy number was essentially stable throughout the timecourse of the experiment (Figures 4.14C and 4.15C).

Thus the LVs tested in these experiments gave similar outcomes in both undifferentiated and differentiated F9 and P19 cells. As we have seen previously with other candidate UCOEs, the rate of reduction in eGFP expression with the novel candidate 455UCOE, 527UCOE and 945UCOE LVs is greater upon differentiation (Figures 4.14A and 4.15A) than in undifferentiated (Figures 4.12A and 4.13A) cells.

A)


EGFP-Positive Cells (%)
60
50
40
30
20
10
0
Day 3
Day 10
Day 17
Day 24
Day 31
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
455UCOE
527UCOE
945UCOE

B)


MFI (arbitrary units)
90
80
70
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
455UCOE
527UCOE
945UCOE

C)

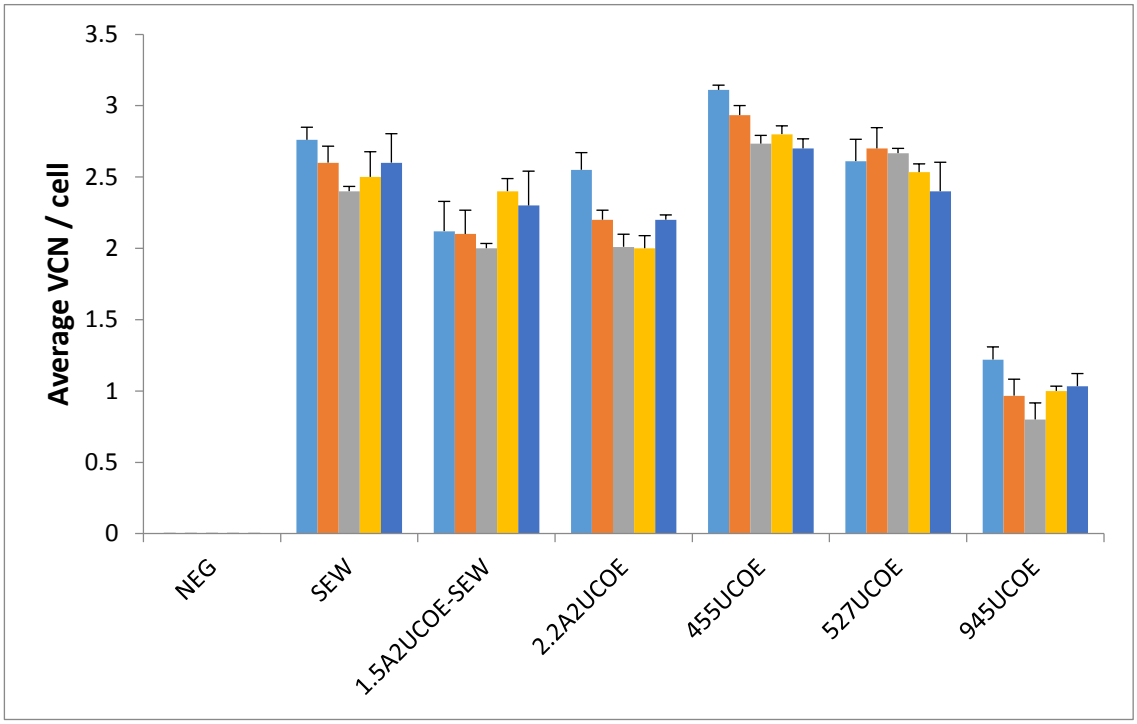
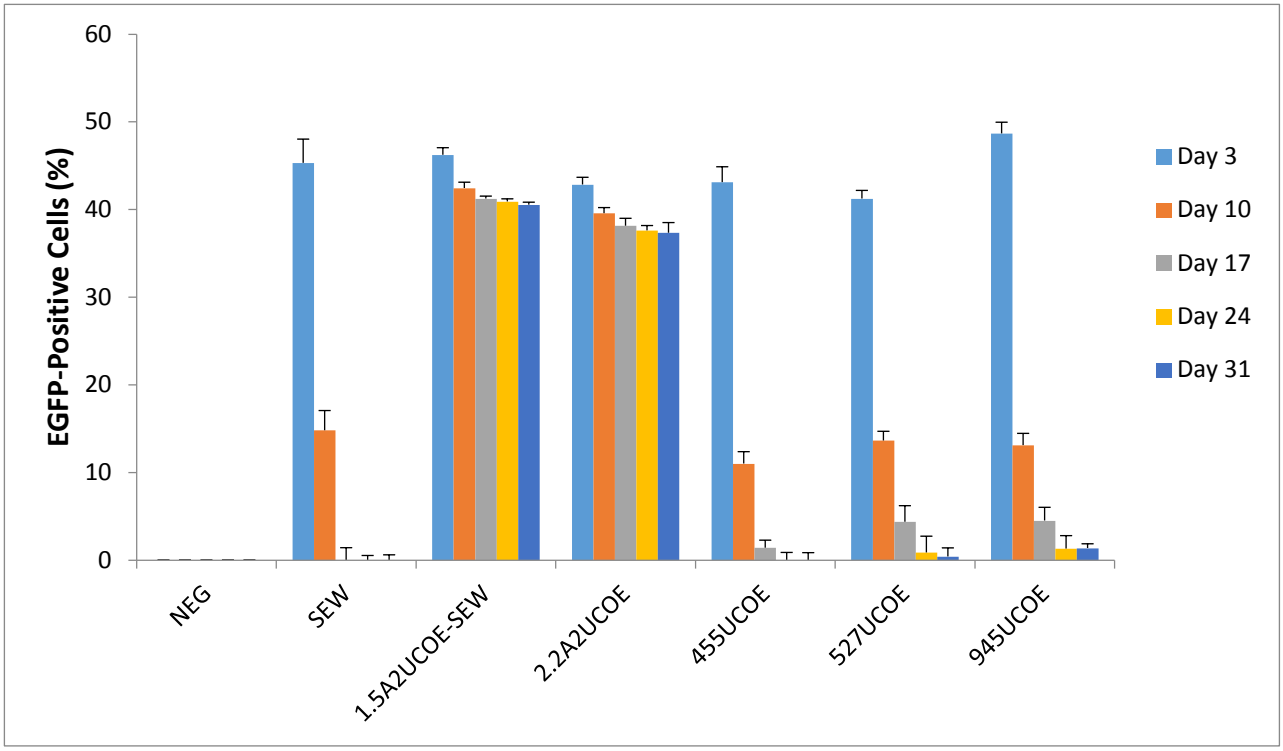
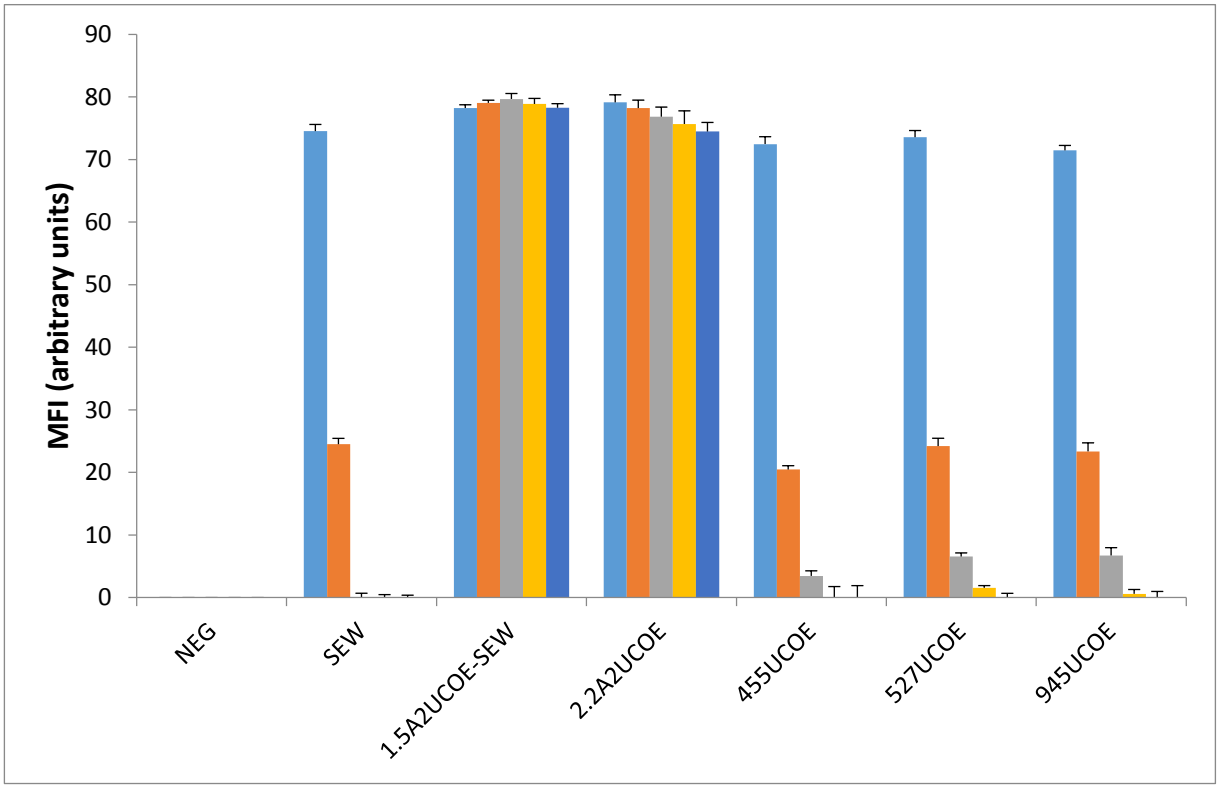


Figure 4.16. Novel candidate 455UCOE, 527UCOE and 945UCOE LVs do not offer protection against silencing in differentiated P19 cells. P19 cells were transduced with novel candidate 455UCOE, 527UCOE and 945UCOE, and control SEW, 1.5A2UCOE-SEW and 2.2A2UCOE LVs (Figure 4.3). Cells were analysed by flow cytometry to detect percentage of eGFP reporter gene expressing (eGFP+) cells, mean fluorescence intensity (MFI) and by RT-Q-PCR for average vector copy number (VCN) per cell. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 31 days post-transduction. (**A**) timecourse of percentage eGFP-positive cells; (Mean + SEM, n=4; **p<0.01). (**B**) As in (A) but showing MFI; (Mean + SEM, n=4; **p<0.01). (**C**) As in (A)/(B) but average VCN/cell; (Mean + SEM, n=4; **p<0.01).

A)


EGFP-Positive Cells (%)
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
455UCOE
527UCOE
945UCOE
Day 3
Day 10
Day 17
Day 24
Day 31

B)


MFI (arbitrary units)
90
80
70
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
455UCOE
527UCOE
945UCOE

C)

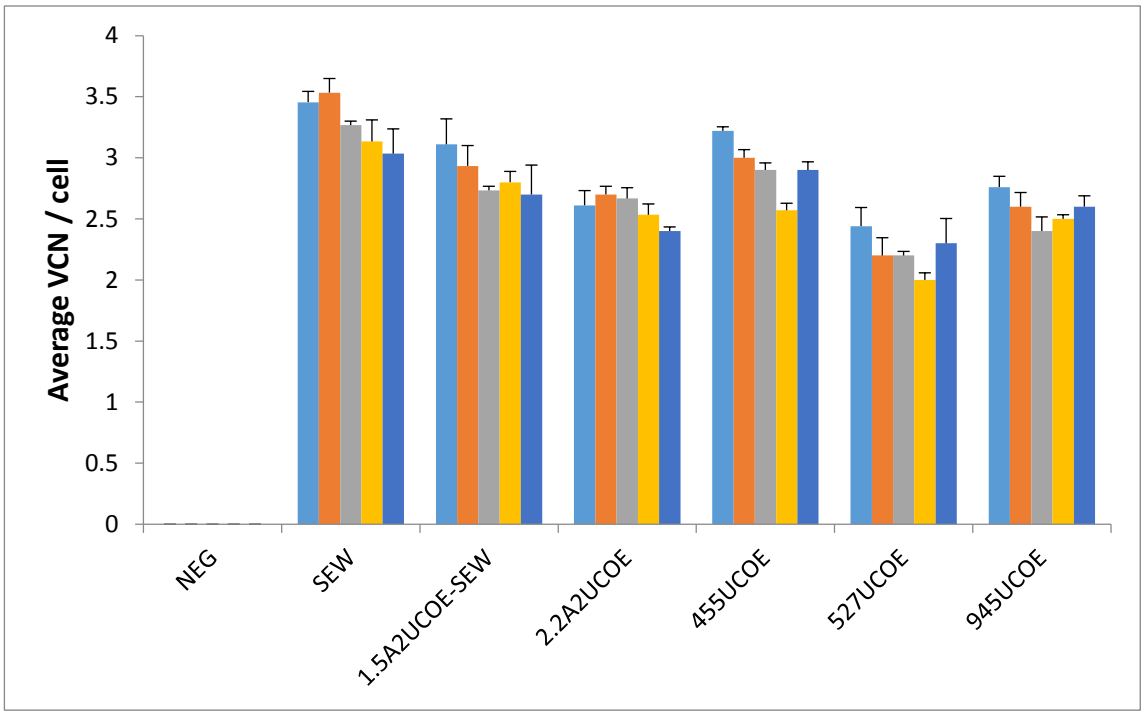


Figure 4.17. Novel candidate 455UCOE, 527UCOE and 945UCOE LVs do not offer protection against silencing in differentiated F9 cells. F9 cells were transduced with novel candidate 455UCOE, 527UCOE and 945UCOE, and control SEW, 1.5A2UCOE-SEW and 2.2A2UCOE LVs (Figure 4.3). Cells were analysed by flow cytometry to detect percentage of eGFP reporter gene expressing (eGFP+) cells, mean fluorescence intensity (MFI) and by RT-Q-PCR for average vector copy number (VCN) per cell. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 31 days post-transduction. (**A**) timecourse of percentage eGFP-positive cells; (Mean + SEM, n=4; **p<0.01). (**B**) As in (A) but showing MFI; (Mean + SEM, n=4; **p<0.01). (**C**) As in (A)/(B) but average VCN/cell; (Mean + SEM, n=4; **p<0.01).

4.11 Functional analysis of candidate 455UCOE, 527UCOE and 945UCOE LVs in differentiated P19 and F9 cells by immunofluorescence staining

As before, in order to confirm the results obtained by flow cytometry (Figures 4.16 and 4.17), P19 and F9 cells which had undergone neuroectodermal and endodermal differentiation, respectively, were stained for appropriate markers and scored under immunofluorescence microscopy. The results of transduced GFP plus β-tubulin III double-positive P19 cells (Figure 4.18) and GFP plus Oct3-4 double-positive F9 cells (Figure 4.19) are in complete agreement with the flow cytometry data (Figures 4.16 and 4.17) and show that the candidate 455UCOE, 527UCOE and 945UCOE constructs fail to protect the linked SFFV promoter from silencing over the course of differentiation.

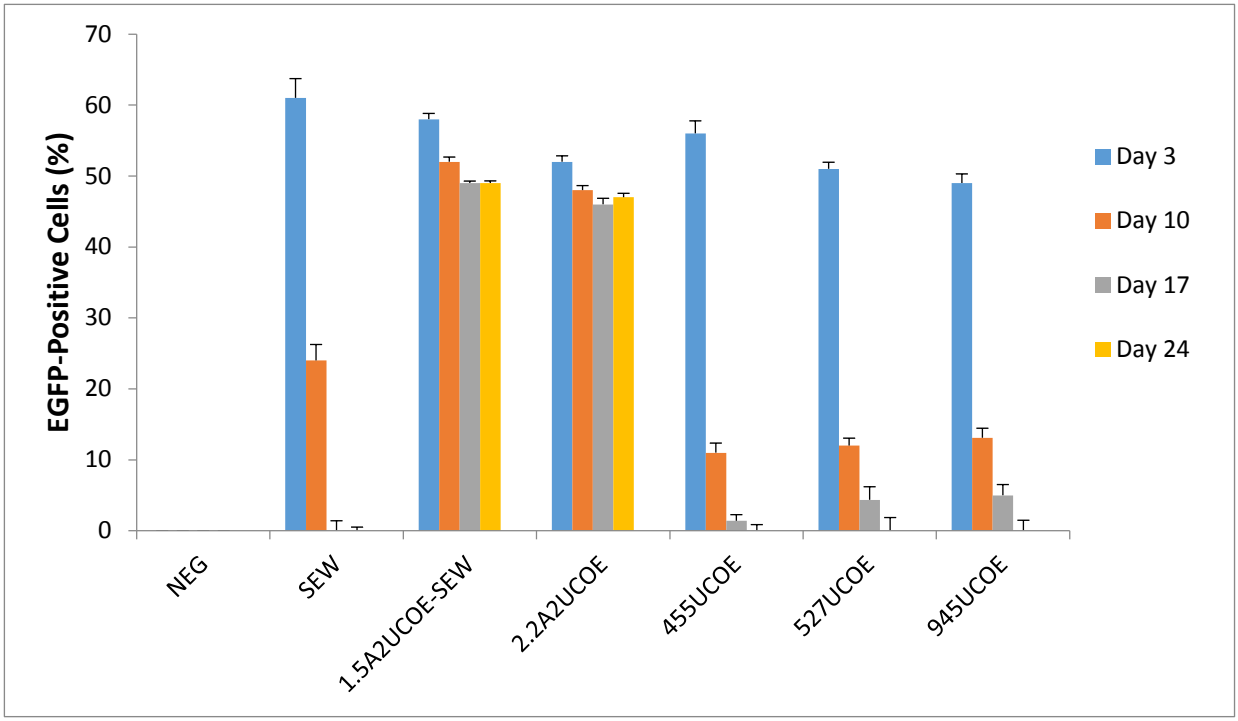


Figure 4.18: Novel candidate UCOEs offer only partial protection against silencing in differentiated P19 cells. P19 cells were transduced with novel candidate 455UCOE, 527UCOE and 945UCOE, and control SEW, 1.5A2UCOE-SEW and 2.2A2UCOE LVs (Figure 4.3) and at 3-days post-transduction induced to undergo differentiation down the neuroectodermal lineage. Cells were analysed at various times following differentiation by immunofluorescence microscopy scoring for GFP plus β-tubulin III double-positive cells. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 24 days post-transduction (Mean + SEM, n=4; **p<0.01).

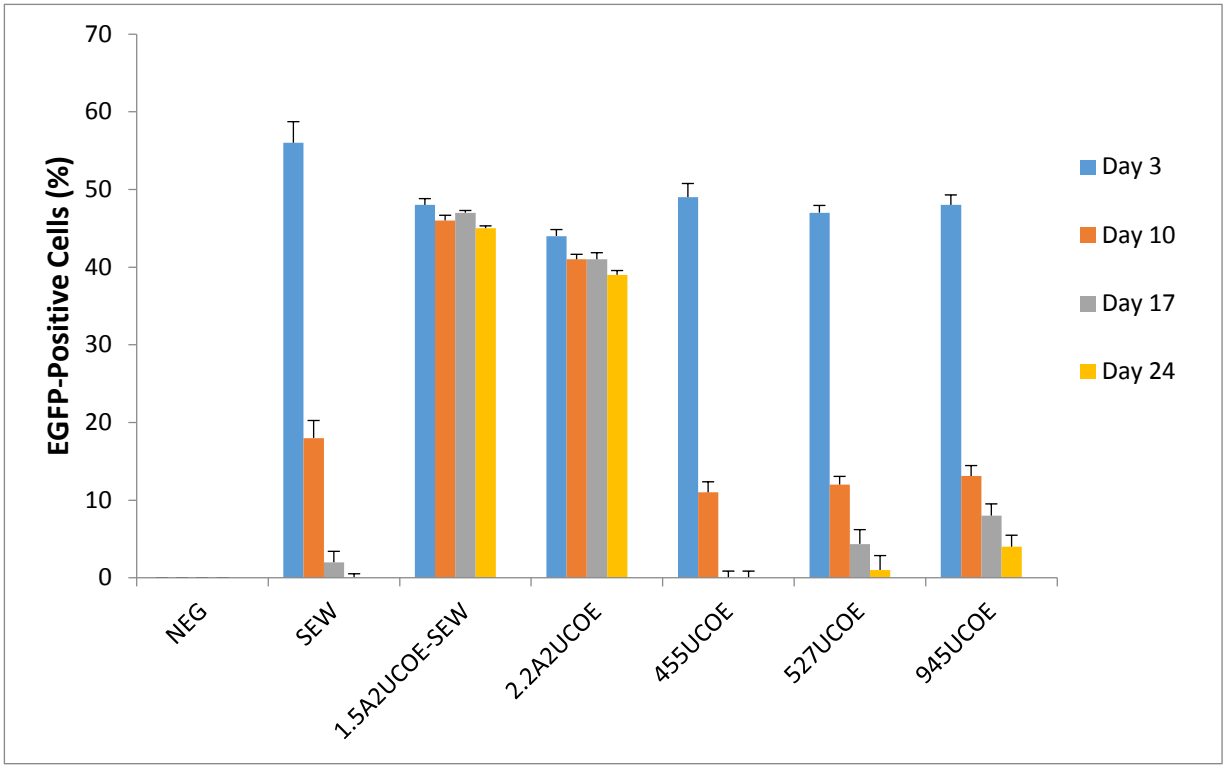


Figure 4.19: Novel candidate UCOEs offer only partial protection against silencing in differentiated F9 cells. F9 cells were transduced with novel candidate 455UCOE, 527UCOE and 945UCOE, and control SEW, 1.5A2UCOE-SEW and 2.2A2UCOE LVs (Figure 4.3) and at 3-days post-transduction induced to undergo differentiation down the endodermal lineage. Cells were analysed at various times following differentiation by immunofluorescence microscopy scoring for GFP plus Oct3-4 double-positive cells. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 24 days post-transduction (Mean + SEM, n=4; **p<0.01).

4.12 Summary and Conclusions

The primary aim of the experiments conducted in this phase of the project was to determine whether sub-regions of the A2UCOE derived from the human *HNRPA2B1-CBX3* locus (Figure 4.1) retain UCOE function. This work was inspired by previous reports, which claimed that (i) a 0.7kb region from the first intron of *CBX3* retains UCOE function when linked to a heterologous CMV promoter (Bandaranayake, Correnti et al. 2011), (ii) promoter-less CpG rich transgenes can establish a methylation-free DNA, open chromatin region in the absence of RNA polymerase II transcriptional activity (Thomson, Skene et al. 2010) and (iii) sub-fragments of CpG islands associated with developmentally regulated genes, including some lacking a promoter, retain their ability to maintain appropriate DNA methylation status in a transgene context (Lienert, Wirbelauer et al. 2011). Some of these observations clearly call into question the dual divergent transcription model of UCOE function (Allen and Antoniou 2007).

The LV constructs Daedalus-F and Daedalus-R (Figure 4.3) were built to directly test the claim that a 0.7kb fragment from the first intron of *CBX3* (Figure 4.1) can confer UCOE activity. Our results from transduced P19 and F9 cells both before and after differentiation down the neuroectodermal and endodermal lineages, respectively, clearly show that this element does not confer stability of expression on a linked SFFV promoter (Figures 4.4 - 4.7). Evidently, the reason for the apparent discrepancy between our results and published data appears to be due to the fact that previous claims of a UCOE capability for this 0.7kb *CBX3* intronic fragment stem from transductions that gave rise to an average LV copy number per cell of 8-10

(Bandaranayake, Correnti et al. 2011). In this scenario only a single integrated vector copy remaining active in a given cell would allow it to score positive for GFP-reporter gene expression and thus give the impression of functional stability. Indeed, the large (55%) drop in MFI observed in cells at 4 weeks post-transduction with the 0.7UCOE-GFP ("Daedalus") LV implies a large degree of vector silencing despite the apparent maintenance in numbers of GFP+ cells (Figure 4.2; (Bandaranayake, Correnti et al. 2011). Our results confirm that in all likelihood silencing of 0.7UCOE (Daedalus) vector expression is taking place. This is particularly evident in our experiments since they were designed to generate cells with a low (1-4) average vector copy number per cell (Figures 4.4C – 4.7C). In addition, the P19 and F9 embryonal carcinoma cell lines we employ may represent a more sensitive and accurate functional gene silencing assay system (Zhang, Thornhill et al. 2007) (Zhang, Frost et al. 2010) than the CHO cells used previously (Bandaranayake, Correnti et al. 2011).

The candidate 0.9UCOE-F and 0.9UCOE-R UCOEs were constructed based on a bioinformatics search of the A2UCOE region for certain (Sp1, CTCF, USF) transcription factor binding sites in an effort to identify a possible MDR (Lienert, Wirbelauer et al. 2011). The linking of this 0.9kb A2UCOE core region that spans both the *HNRPA2B1* and *CBX3* promoter/transcriptional start sites (Figure 4.1) to the SFFV promoter in the SEW vector (Figure 4.3), resulted in only a partial protection against silencing in both undifferentiated and differentiated P19 (Figures 4.8, 4.10, 4.12) and F9 (Figures 4.9, 4.11, 4.13) cells. Thus, although the 0.9kb core A2UCOE promoter fragment may constitute an MDR, it does not retain a full UCOE capability unlike the larger 1.5kb A2UCOE, which extends further at both the *HNRPA2B1* and

CBX3 ends. This difference in UCOE activity between these two A2UCOE core elements may be due the reduction in the length of the CpG island beyond a crucial point within the 0.9kb fragment such that it is unable to stably maintain its normally inherent open chromatin structure.

The construction of the candidate 455UCOE, 527UCOE and 945UCOE LVs was based on the observation that CpG-rich DNA fragments devoid of promoter activity can establish an open chromatin structure in a transgene context (Thomson, Skene et al. 2010). The 455UCOE, 527UCOE and 945UCOE LVs contain fragments from the region immediately downstream of the two alternative first exons of *CBX3*, which possess the highest CpG density and also lack a promoter (Figure 4.1). It was therefore hypothesised that these 455bp, 527bp and 945bp fragments may be able to establish an open chromatin structure independent of promoter activity as previously suggested (Thomson, Skene et al. 2010) and confer stability on a linked SFFV promoter (Figure 4.3). Our results in both undifferentiated and differentiated P19 (Figures 4.14, 4.16, 4.18) and F9 (Figures 4.15, 4.17, 4.19 and 15) cells clearly show that none of these *CBX3*-derived fragments are able to confer stability of expression and thus do not possess a UCOE capability. Indeed, the rate of reduction in expression was comparable between the SEW vector that contains the SFFV promoter alone, and those LVs with the linked 455UCOE, 527UCOE and 945UCOE elements (Figures 4.14A to 4.15A) demonstrating no UCOE activity for the 455bp and 527bp fragments either alone or in combination as the 945bp element.

In conclusion, none of the novel candidate UCOE sub-fragments of the A2UCOE region either with (0.9UCOE-F, 0.9UCOE-R) or without (Daedalus-F, Daedalus-R; 455UCOE, 527UCOE, 945UCOE) associated promoter activity retained a full UCOE capability. Only the 0.9UCOE-F and especially 0.9UCOE-R constructs that encompass the promoters and transcriptional start sites of *HNRPA2B1* and *CBX3* possessed a significant but still partial ability to negate transgene silencing. Thus, despite published evidence to the contrary we were unable by these targeted experiments to identify an A2UCOE sub-region devoid of promoter activity that still retained a full UCOE function.

Chapter 5
Results

Minimising the A2UCOE for direct transgene expression from the innate *HNRPA2B1* promoter

Chapter 5

Results: Minimising the A2UCOE for direct transgene expression from the innate *HNRPA2B1* promoter

5.1 Introduction and background

It has previously been shown that the innate *HNRPA2B1* promoter of the 2.2kb A2UCOE (2.2A2UCOE; Figure 5.1) can provide stable transgene expression from both plasmid (Antoniou, Harland et al. 2003) and lentiviral (Zhang, Thornhill et al. 2007) vectors. (See also, for example, Figures 4.8 and 4.9). As we failed to identify a CpG-rich A2UCOE sub-fragment either with or without associated promoter activity that fully retained UCOE function (Chapter 4), we next sought to determine if the 2.2A2UCOE could be reduced in size but still retain full UCOE activity. In a previously published study it was shown that a 0.9kb A2UCOE (0.9A2UCOE; Figure 5.1) again where transcription was directly from the *HNRPA2B1*promoter, did not provide stable expression in P19 cells transduced with an LV containing an 0.9A2UCOE-eGFP cassette (Knight, Zhang et al. 2012). We, therefore, constructed LVs containing an A2UCOE-eGFP transgene deletion series between 2.2kb and 0.9kb in length (Figure 5.1). These vectors were tested for UCOE function in our standard murine embryonal carcinoma P19 and F9 cell assay system. The stability, potency, and reproducibility of expression of the eGFP transgene were analyzed in both differentiated and undifferentiated P19 and F9 cells.

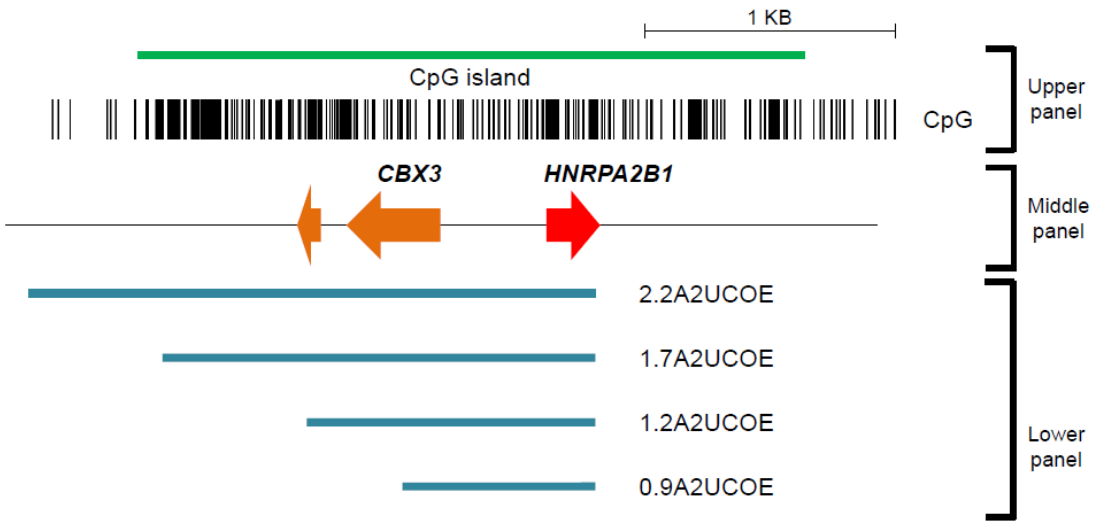


Figure 5.1. Relative positions of the 2.2A2UCOE and sub-fragments analysed for UCOE activity. Upper panel: CpG density map and methylation-free CpG island. Middle panel: Positions of the two alternative first exons of *CBX3* (green arrows) and the first exon of *HNRPA2B1* (red arrow). Lower panel: positions of A2UCOE sub-region fragments analysed for UCOE activity. 2.2A2UCOE: 2.2kb *Bam*HI-*Tth*111I element. 1.7A2UCOE: 5' deletion of the 2.2kb element to a *Hpa*I site within the first intron of *CBX3* to give a 1.7kb fragment. 1.2A2UCOE: further 5' deletion of 2.2A2UCOE element to a *Bsm*BI site at the end of the alternative second exon of *CBX3* to give a 1.2kb fragment. 0.9UCOE: 5'deletion sub-fragment of the 2.2A2UCOE to a *Csp*45I site at the end of the alternative first exon of *CBX3* to give a 0.9kb fragment. Note: all fragments terminate at a *Tth*111I site 15bp from the end of exon I of *HNRPA2B1*.

5.2 LV used in this study

The LV constructs SEW (SFFV-eGFP-WPRE) and 1.5A2UCOE-SEW are as previously described in Figures 3.1 and 4.3. The 2.2A2UCOE, 1.7A2UCOE, 1.2A2UCOE and 0.9A2UCOE vectors are illustrated in Figure 5.2.

Note: the 1.7A2UCOE and 1.2A2UCOE LVs were generated by Gautom Baruah and 0.9A2UCOE vector was constructed by Thomas Broughton.

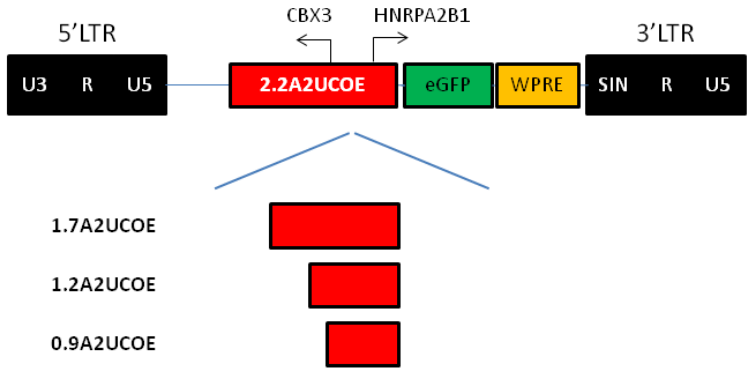


Figure 5.2. Illustration of the 2.2A2UCOE deletion series lentiviral vectors.

Top panel: A standard self-inactivating (SIN) lentiviral vector backbone containing a 2.2kb A2UCOE (2.2A2UCOE) fragment with the innate *HNRPA2B1* promoter driving expression of an enhanced green fluorescent protein (eGFP) reporter gene with downstream woodchuck hepatitis virus post-transcriptional regulatory element (WPRE). Lower panel: 5' (*CBX3* end) deletion series of the 2.2A2UCOE consisting of 1.7kb, 1.2kb and 0.9kb sub-fragments. LTR: long terminal repeat.

5.2.1. 2.2A2UCOE

The 2.2kb A2UCOE-eGFP cassette was generated by subcloning of the 2.2kb *Bam*HI to *Tth*111I A2UCOE fragment (Figure 5.1; 2.2A2UCOE; Williams *et al*., 2005) in place of the SFFV promoter in the SEW LV backbone (Figure 5.2, top panel). The orientation of the 2.2A2UCOE is such that the *HNRPA2B1* promoter drives expression of the linked eGFP reporter gene. This 2.2A2UCOE LV has previously been shown to provide stable expression in P19 cell assays (Zhang, Frost et al. 2010).

5.2.2. 1.7A2UCOE and 1.2A2UCOE

The 1.7UCOE and 1.2UCOE constructs were generated by deleting the 2.2A2UCOE to *Hpa*I and *Bsm*BI sites, respectively, at the *CBX3* end of this element to give 1.7kb and 1.2kb sub-fragments (Figure 5.1). Again these are joined directly via the *HNRPA2B1* promoter to an eGFP reporter gene to produce the desired LVs (Figure 5.2)

5.2.3. 0.9A2UCOE

The 0.9A2UCOE construct was generated by a further deletion of the 2.2A2UCOE element to a *Csp*45I site within the alternative first exon of *CBX3* (Figure 5.1) and linked to an eGFP reporter gene (Figure 5.2).

5.3 Lentiviral vector titration in HEK293T cells

Cells were collected for analysis by flow cytometry three days post-transduction of HEK293T cells with serial dilutions of viral stocks collected at both 2 and 3 days post-transfection from each production run for all LVs: SEW (SFFV-eGFP-WPRE), 1.5A2UCOE-SEW, 2.2A2UCOE, 1.7A2UCOE and 1.2A2UCOE (Table 5.1).

LV	Volume (µL) of viral vector stock			
	2 µL	0.2 µL	0.02 µL	0.002 µL
	Percentage eGFP+ cells 1st / 2nd viral harvest			
SEW	72.45 / 61.43	12.23 / 10.12	2.5 / 1.22	0.54 / 0.23
1.5A2UCOE-SEW	78.48 / 66.2	16.8 / 11.18	3.5 / 2.14	0.82 / 0.51
2.2A2UCOE	73.6 / 68.16	14.16 / 11.22	3.04 / 1.33	0.72 / 0.42
1.7A2UCOE	71.1 / 63.44	12.05 / 8.12	2.04 / 1.02	0.6 / 0.12
1.2A2UCOE	84.18 / 72.88	18.11 / 10.4	4.3 / 1.86	0.48 / 0.16

Table 5.1. Titration of lentiviral vector (LV) preparations. A $2x10^5$ aliquots of HEK293T cells were transduced with 0.002-2 µL of a given LV stock from the first and second harvest of virus from culture supernatant during the time of preparation. Analysis of eGFP-positive cells at 3-days post-transduction was by flow cytometry. LV constructs are as described in Figure 5.2.

5.3.1 Calculation of LV titre

Viral titre (Table 5.2) was calculated based on the dilution of each harvest that gave a percentage of eGFP-positive cells between 1% and 10%. Pools of cells with percentages of eGFP-positive cells higher than 10% are very likely to contain multiple integrations of the vector, while very low scores could be false positives.

On day one, $2x10^5$ 293T cells were seeded in each 6-well. Thus, the percentage of eGFP-positive cells from flow cytometry analysis, reflects the percentage of the initial cell population that was successfully transduced and hence the number of infectious units added in the specified well. The dilution factor or the volume of the lentivirus preparation that was used is known, so the calculation of the number of infectious units per ml from that point onwards is straightforward.

LV	Titre (iu/mL)	
	1st harvest	2nd harvest
SEW	2.42 x 10^8	1.66x 10^8
1.5A2UCOE-SEW	2.85 x 10^8	1.96 x 10^8
2.2A2UCOE	2.35 x 10^8	1.28 x 10^8
1.2A2UCOE	2.58 x 10^8	1.24x 10^8
1.7A2UCOE	2.62 x 10^8	1.38 x 10^8

Table 5.2: Lentiviral vector (LV) titres. Viral titre as infectious units per ml (iu/mL) was calculated based on the dilution of 1st and 2nd harvest of virus that gave a percentage of eGFP-positive HEK293T cells between 1% and 10% (Table 3.1). LV constructs are as described in Figure 3.1.

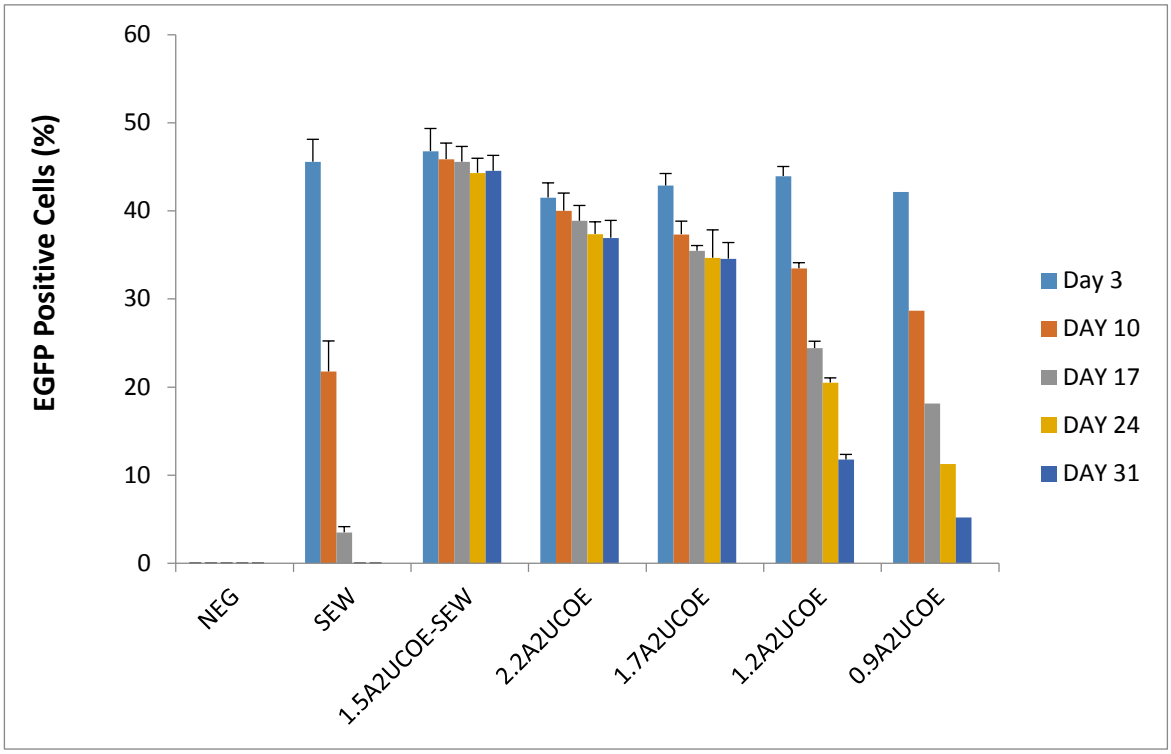
5.4 Functional assay of candidate UCOEs 1.2A2UCOE and 1.7A2UCOE in undifferentiated P19 and F9 cells

P19 and F9 cells were transduced with the generated LVs at an MOI of 3 or 6, with the intention to start the experiment within the range of 40-60% eGFP-positive cells in all pools. The transduced cells and negative control were then propagated in culture and assayed as a time course by flow cytometry, to determine the percentage of eGFP-positive cells and the mean fluorescence intensity (MFI) for each repeat transduction for each vector. Cell cultures were analysed every 7 days from day 3 post-transduction and extending to 45 days. Cell samples were taken at regular intervals for the flow cytometry analysis, to monitor the expression levels of the eGFP reporter gene. In addition, DNA was extracted from cells at each time point for RT-qPCR in order to determine the average vector copy number per cell.

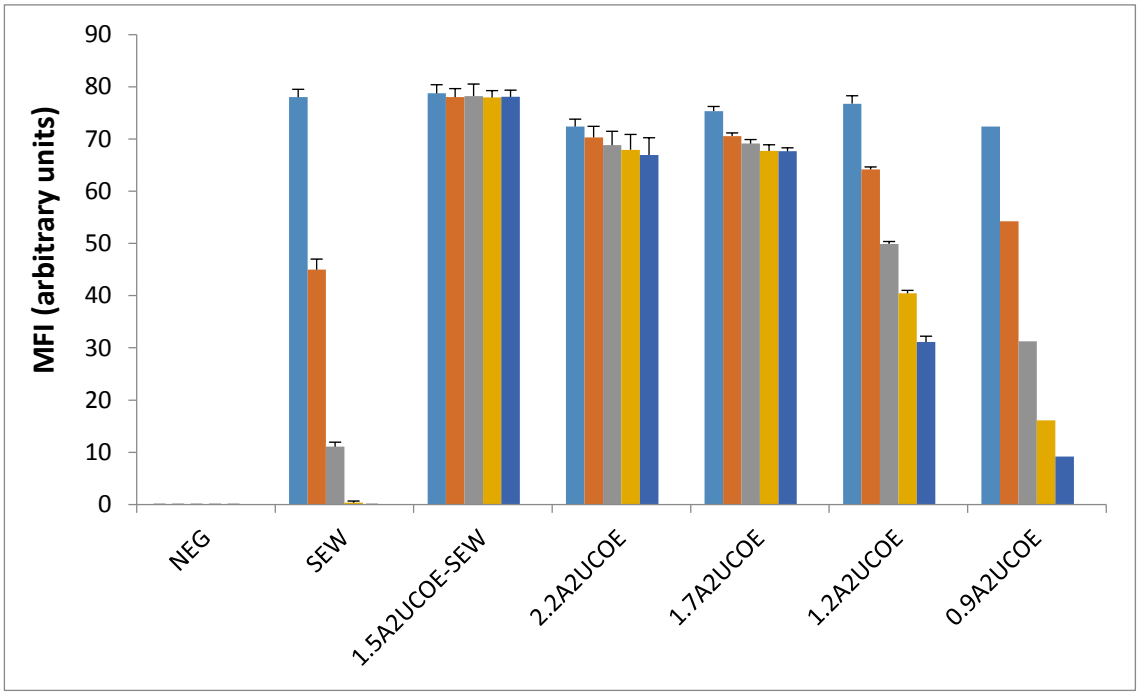
Figure 5.2A shows the flow cytometry time course results depicting percentage of eGFP+ cells. Although similar initial transduction efficiency was obtained with all vectors (45-60% eGFP+ cells), expression from the SEW (SFFV-eGFP-WPRE), rapidly declined from 47% to 4% positive cells within 17 days. In contrast, the proportion of eGFP+ cells from the 1.5A2UCOE-SEW and 2.2kbA2UCOE vectors remained completely stable over the 31-day period of culture. Therefore, whilst differences in the percentage of eGFP-positive cells transduced with 1.5A2UCOE-SEW and 2.2kbA2UCOE were significantly higher when compared with the results obtained for SEW and 1.2A2UCOE, the decrease of eGFP expression of the vector 1.2A2UCOE is parallel when compared with SEW. The expression of eGFP in SEW transduced cells had

dropped almost 80% after 2 weeks post-transduction. Similarly, with the exception of 1.5A2UCOE-SEW and 2.2kbA2UCOE, the reduction observed in eGFP-positive cells was 30% to 40% for 1.2A2UCOE over the same time period. However, the expression of eGFP in 1.7A2UCOE has remained stable over the same time period. The values of mean fluorescent intensity (MFI) for all vectors paralleled the eGFP expression results (Figure 5.2B). MFI was stable in the 1.5A2UCOE-SEW, 2.2kbA2UCOE and 1.7A2UCOE transduced cells, and unstable in the case of SEW and 1.2A2UCOE.

A)


EGFP Positive Cells (%)
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
1.7A2UCOE
1.2A2UCOE
0.9A2UCOE
Day 3
DAY 10
DAY 17
DAY 24
DAY 31

B)


MFI (arbitrary units)
90
80
70
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
1.7A2UCOE
1.2A2UCOE
0.9A2UCOE

C)

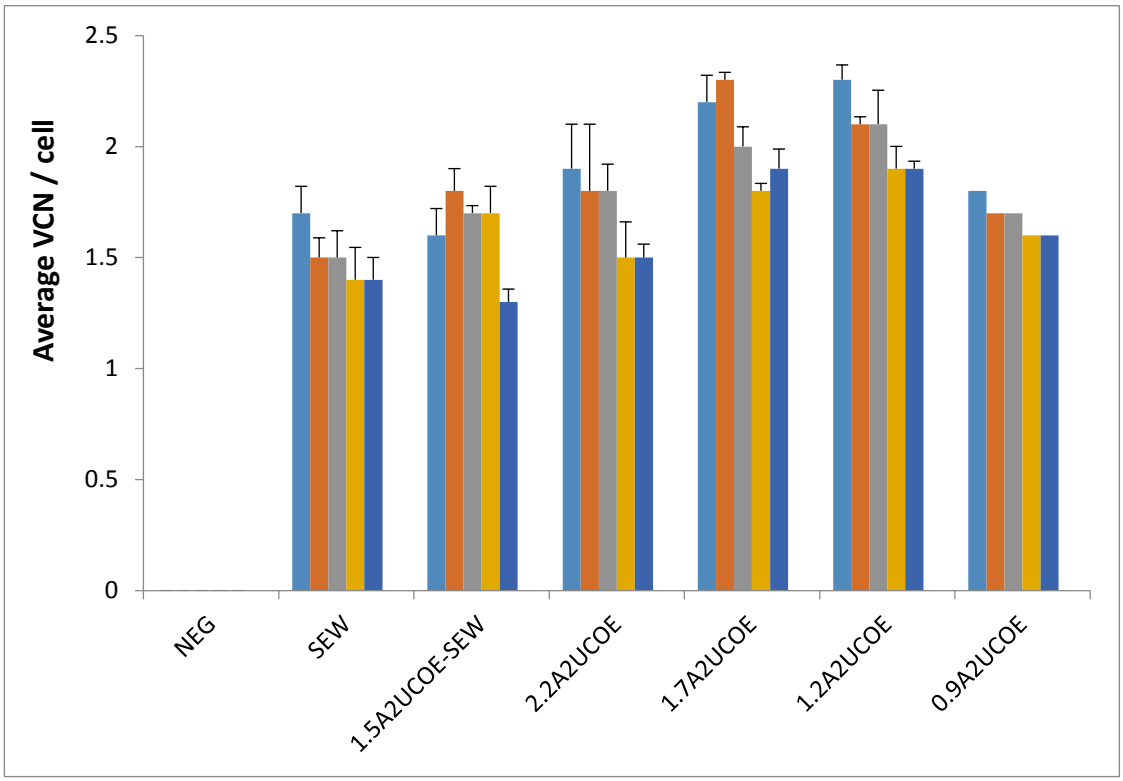
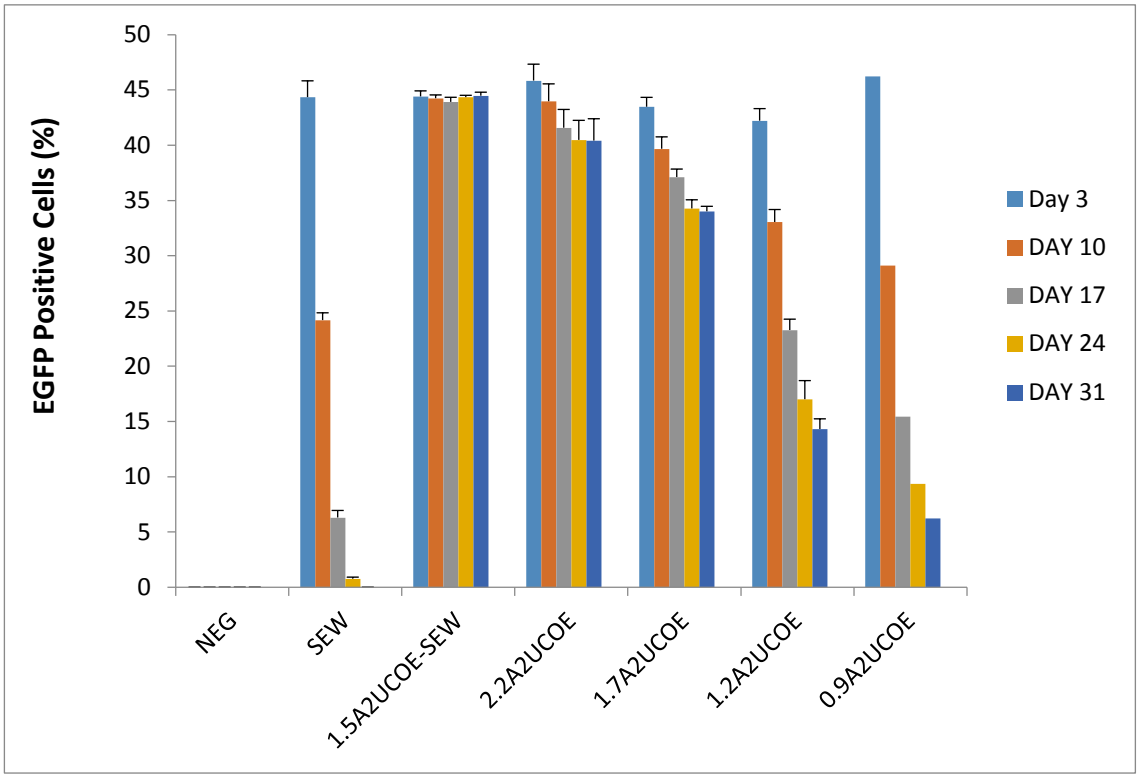


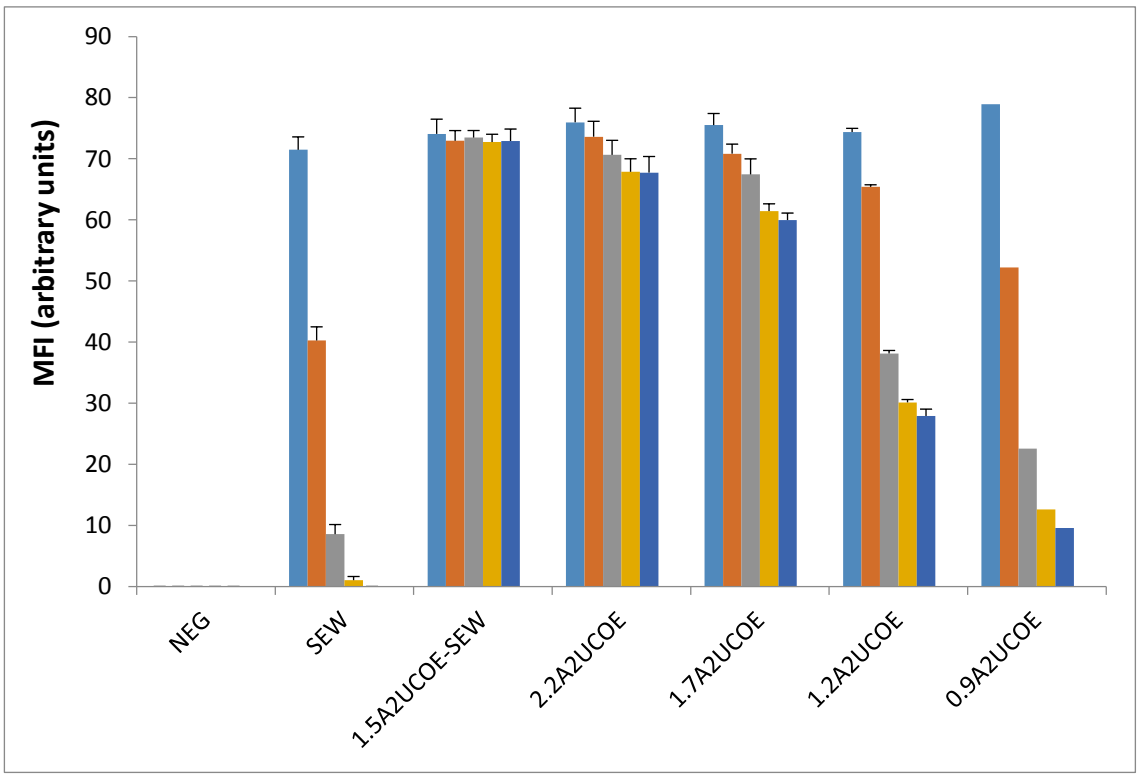
Figure 5.2: Novel candidate UCOEs offer only partial protection against silencing in undifferentiated P19 cells. P19 cells were transduced with novel candidate UCOE 1.2A2UCOE, 1.7A2UCOE and control SEW, 2.2kbA2UCOE and 1.5A2UCOE-SEW lentiviral vectors (Figure 3.1). Cells were analysed by flow cytometry to detect percentage of eGFP reporter gene expressing (eGFP+) cells, mean fluorescence intensity (MFI) and by RT-Q-PCR for average vector copy number (VCN) per cell. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 31 days post-transduction. (**A**) timecourse of percentage eGFP-positive cells; (Mean + SEM, n=4; **p<0.01). (**B**) As in (A) but showing MFI; (Mean + SEM, n=4; **p<0.01). (**C**) As in (A)/(B) but average VCN/cell; (Mean + SEM, n=4; **p<0.01).

Figure 5.3A shows the flow cytometry time course results depicting percentage of eGFP+ cells in undifferentiated F9 cells. The results show a parallel activity as with the P19 cells above. Although similar initial transduction efficiency was obtained with all vectors (39–50% eGFP+ cells), expression from the SEW (SFFV-eGFP-WPRE), rapidly declined from 45% to 5% positive cells within 17 days. In contrast, the proportion of eGFP+ cells from the 1.5A2UCOE-SEW, 2.2kbA2UCOE and 1.7A2UCOE vectors remained completely stable over the 31-day period of culture. Therefore, whilst differences in the percentage eGFP-positive cells transduced with 1.5A2UCOE-SEW, 2.2kbA2UCOE and 1.7A2UCOE were significantly higher when compared with the results obtained for SEW and 1.2kb vectors, the decrease of eGFP expression of the vector 1.2A2UCOE is slower when compared with SEW. The expression of eGFP in SEW transduced cells had dropped almost 80% after 2 weeks post-transduction. The values of mean fluorescent intensity (MFI) for all vectors paralleled the eGFP expression results (Figure 5.3B). MFI was stable in the 1.5A2UCOE-SEW, 2.2kbA2UCOE and 1.7A2UCOE transduced cells, unstable in the case of SEW and partially stable in 1.2A2UCOE.

A)


EGFP Positive Cells (%)
50
45
40
35
30
25
20
15
10
5
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
1.7A2UCOE
1.2A2UCOE
0.9A2UCOE
Day 3
DAY 10
DAY 17
DAY 24
DAY 31

B)


MFI (arbitrary units)
90
80
70
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
1.7A2UCOE
1.2A2UCOE
0.9A2UCOE

C)

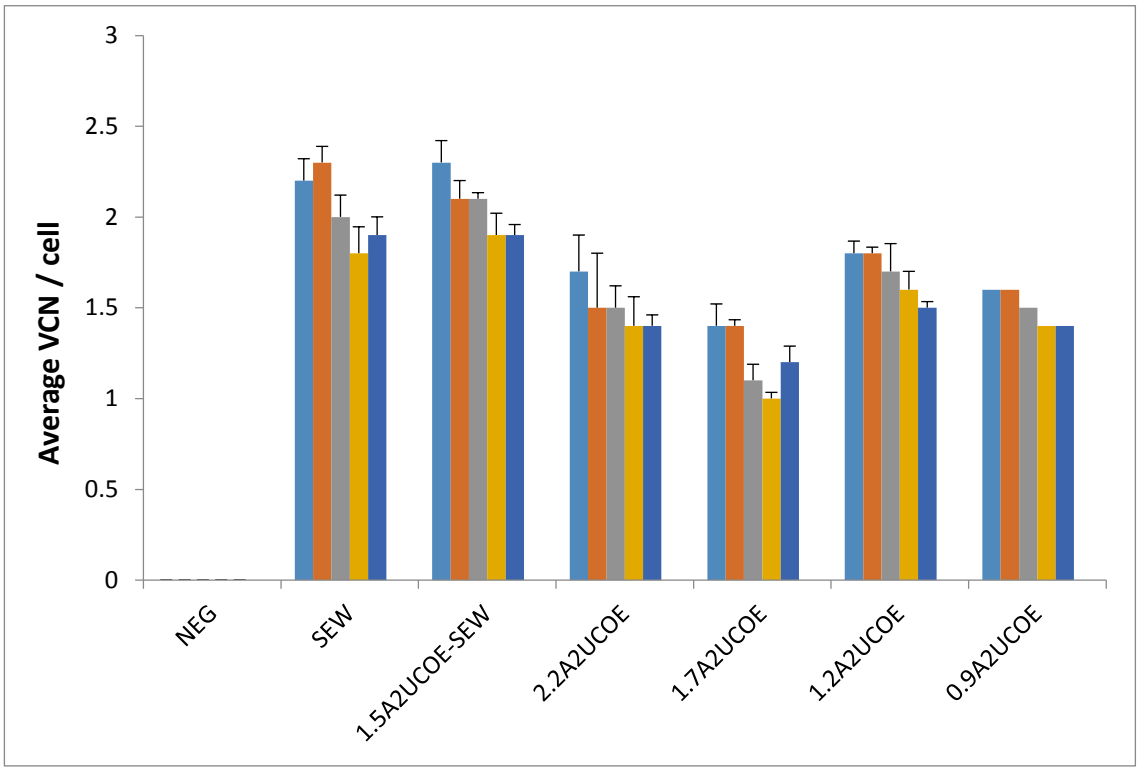


Figure 5.3: Novel candidate UCOEs offer only partial protection against silencing in undifferentiated F9 cells. F9 cells were transduced with novel candidate UCOE 1.2A2UCOE, 1.7A2UCOE and control SEW, 2.2kbA2UCOE and 1.5A2UCOE-SEW lentiviral vectors (Figure 5.1). Cells were analysed by flow cytometry to detect percentage of eGFP reporter gene expressing (eGFP+) cells, mean fluorescence intensity (MFI) and by RT-Q-PCR for average vector copy number (VCN) per cell. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 31 days post-transduction. (**A**) timecourse of percentage eGFP-positive cells; (Mean + SEM, n=4; **p<0.01). (**B**) As in (A) but showing MFI; (Mean + SEM, n=4; **p<0.01). (**C**) As in (A)/(B) but average VCN/cell; (Mean + SEM, n=4; **p<0.01).

5.5 Functional assay of candidate UCOEs 1.2A2UCOE, 1.7A2UCOE in differentiated P19 and F9 cells

The novel candidate UCOE vectors were able to confer partial stability of expression in undifferentiated P19 (Figure 5.4) and F9 (Figure 5.5) cells. We next evaluated the ability of these novel LVs to stabilize expression upon differentiation of P19 and F9 cells down the neuroectodermal and endodermal lineages respectively.

Cells transduced with each LV were first induced to form embryoid bodies in retinoic acid-containing differentiation medium via culture on non-adhesive plastic, which is a necessary pre-requisite state for later neuronal and endodermal differentiation. Cells subsequently differentiate into a mix of neuroectodermal and parietal endoderm cell types, of which fibroblast-like cells were the predominant type at the initial stage of differentiation (2-3 days after plating for P19 cells and 5-6 days for F9 cells), but with neurons and endoderm cells also present. The cultures were then enriched for non-dividing cells by the application of 10 μM cytosine arabinoside, in which neurons and endoderm cells appear as the only survivors. The optimum incubation time to form embryoid bodies and to differentiate the cells into neuroectodermal and endoderm cell types was determined as 2-3 days for P19 cells and 5-6 days for F9 cells.

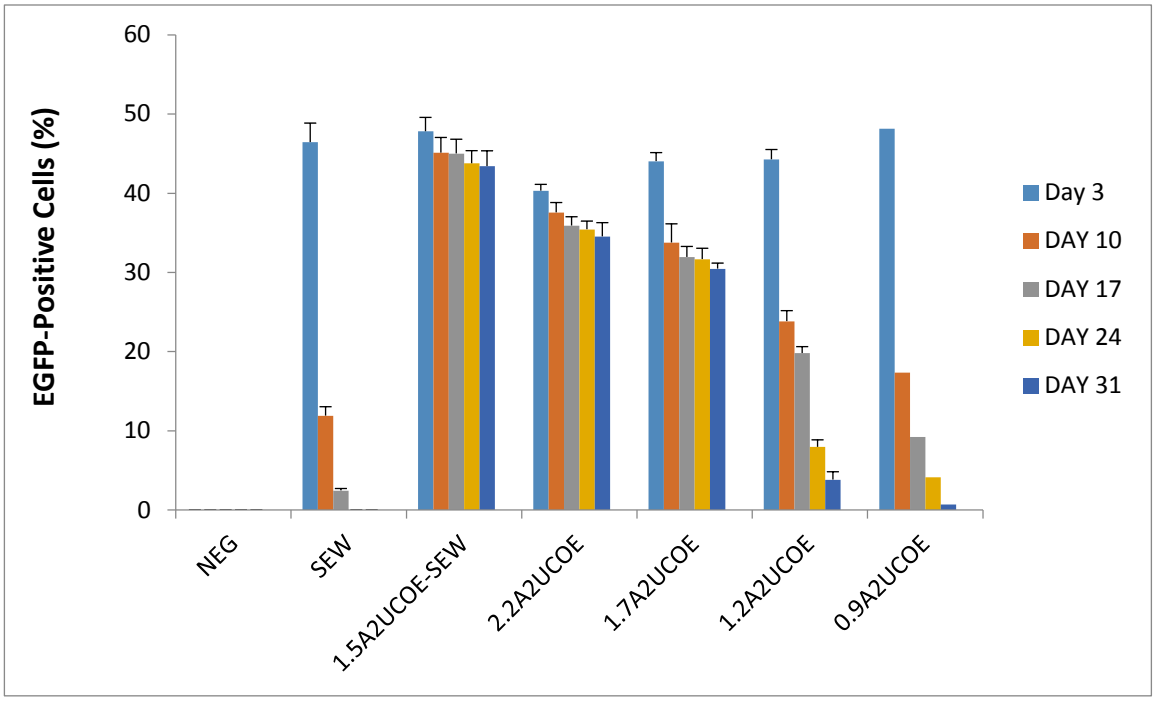
The expression results from triplicate samples of cells transduced with the 1.5A2UCOE-SEW, 2.2A2UCOE, 1.2A2UCOE, 1.7A2UCOE LVs following differentiation is presented in Figures 5.4 and 5.5. Our data from the differentiation experiments show that the control LVs carrying the UCOE construct (1.5A2UCOE-SEW and 2.2A2UCOE) have remained stable after the P19 and F9 cells were

differentiated down the neuronal and endoderm lineages respectively (Figures 5.4A and 5.5A). Also. The 1.7A2UCOE vector has remained stable like the two A2UCOE control vectors (Figures 3.4A and 3.5A). However, the novel candidate UCOE vector 1.2A2UCOE show a gradual reduction in eGFP expression whilst the silencing control SEW vector was rapidly repressed (Figures 5.4A and 5.5A).

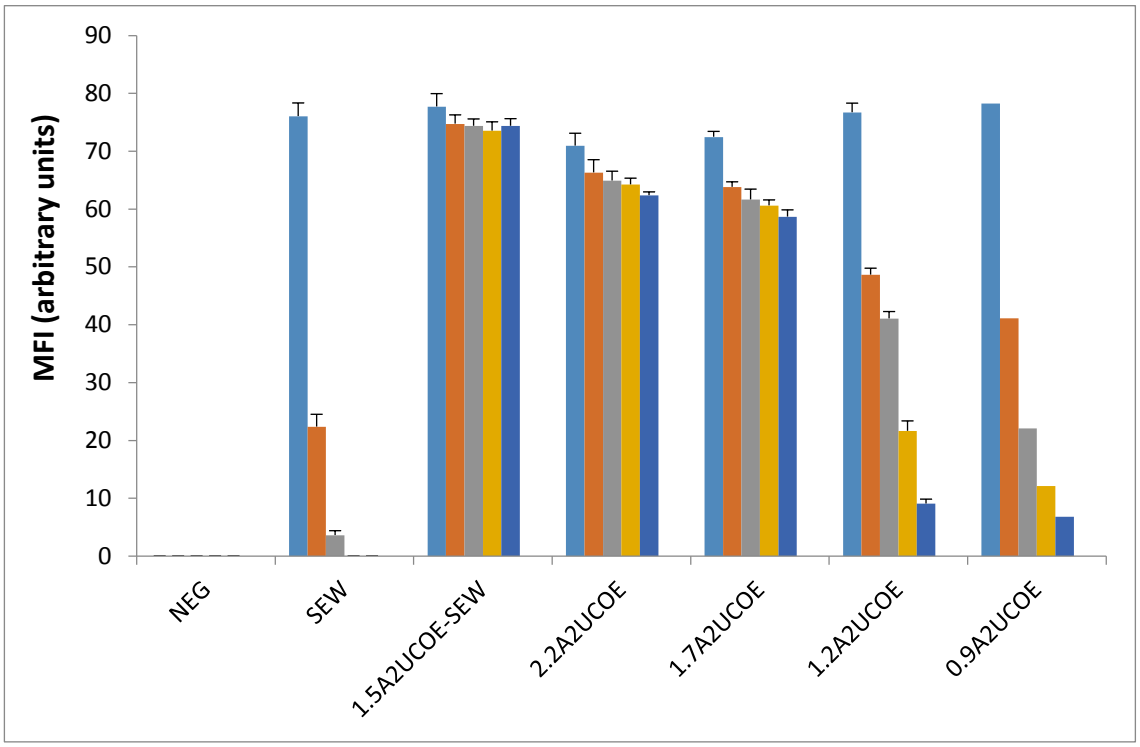
As in the case of undifferentiated cells (Figures 5.2 and 5.3), MFI values upon differentiation of both P19 and F9 cells paralleled the eGFP+ cell expression data (Figures 5.4B and 5.5B) whilst average vector copy number was essentially stable throughout the timecourse of the experiment (Figures 5.4C and 5.5C).

Thus, the LVs tested in these experiments gave similar outcomes in undifferentiated and differentiated F9 and P19 cells. However, it is perhaps noteworthy that the rate of reduction in eGFP expression with the SEW and novel candidate UCOE 1.2A2UCOE LVs is greater upon differentiation (Figures 5.4A and 5.5A) than in undifferentiated (Figures 5.2A and 5.3A) cells.

A)


EGFP-Positive Cells (%)
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
1.7A2UCOE
1.2A2UCOE
0.9A2UCOE
Day 3
DAY 10
DAY 17
DAY 24
DAY 31

B)


MFI (arbitrary units)
90
80
70
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
1.7A2UCOE
1.2A2UCOE
0.9A2UCOE

C)

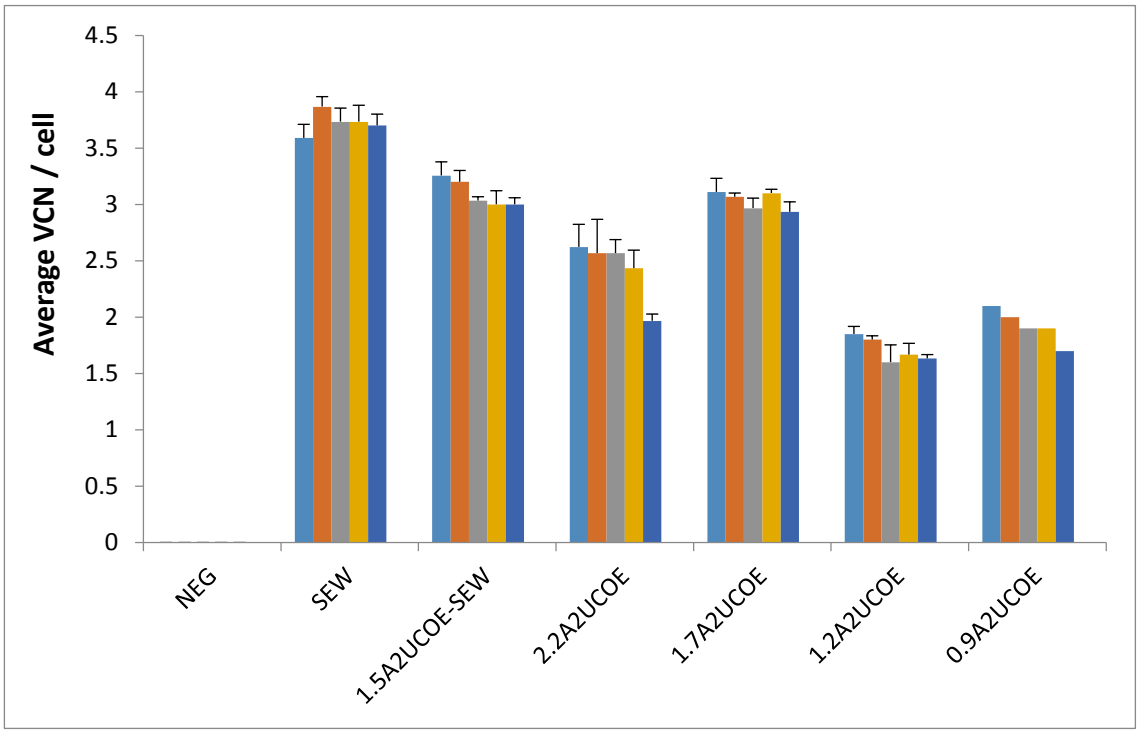
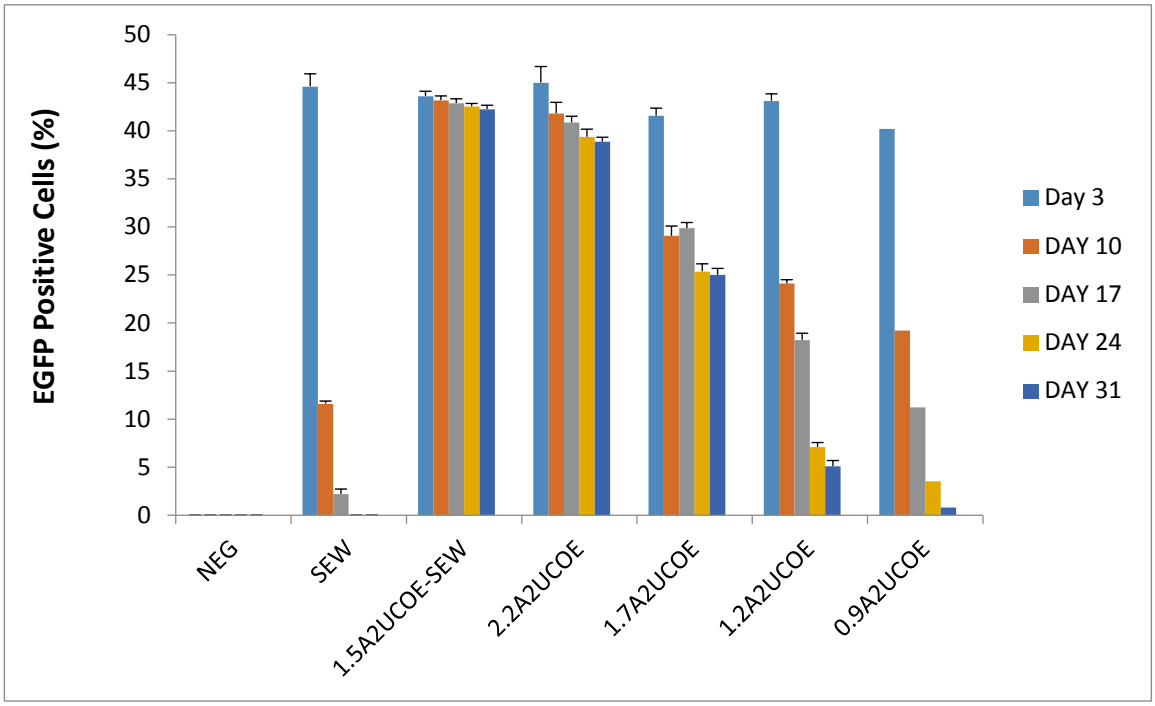
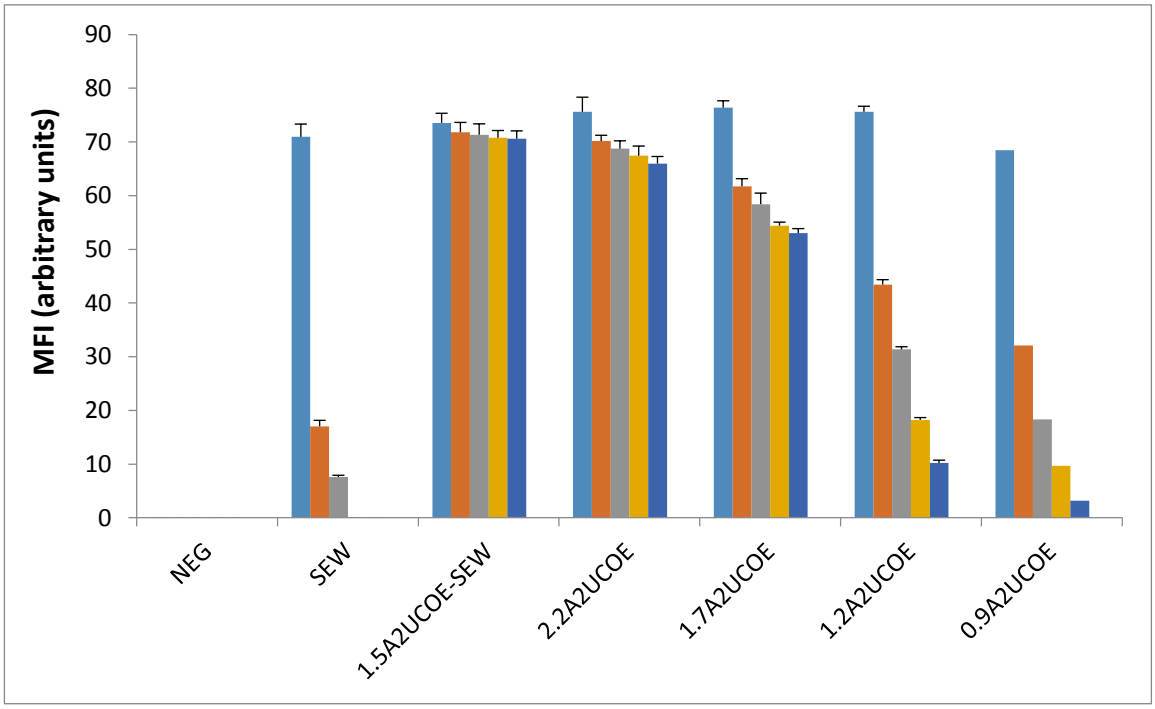


Figure 5.4: Novel candidate UCOEs offer only partial protection against silencing in differentiated P19 cells. P19 cells were transduced with novel candidate UCOE 1.2A2UCOE, 1.7A2UCOE and control SEW, 2.2kbA2UCOE and 1.5A2UCOE-SEW lentiviral vectors (Figure 5.1). Cells were analysed by flow cytometry to detect percentage of eGFP reporter gene expressing (eGFP+) cells, mean fluorescence intensity (MFI) and by RT-Q-PCR for average vector copy number (VCN) per cell. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 45 days post-transduction. (**A**) timecourse of percentage eGFP-positive cells; (Mean + SEM, n=4; **p<0.01). (**B**) As in (A) but showing MFI; (Mean + SEM, n=4; **p<0.01). (**C**) As in (A)/(B) but average VCN/cell; (Mean + SEM, n=4; **p<0.01).

A)


EGFP Positive Cells (%)
50
45
40
35
30
25
20
15
10
5
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
1.7A2UCOE
1.2A2UCOE
0.9A2UCOE
Day 3
DAY 10
DAY 17
DAY 24
DAY 31

B)


MFI (arbitrary units)
90
80
70
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
1.7A2UCOE
1.2A2UCOE
0.9A2UCOE

C)

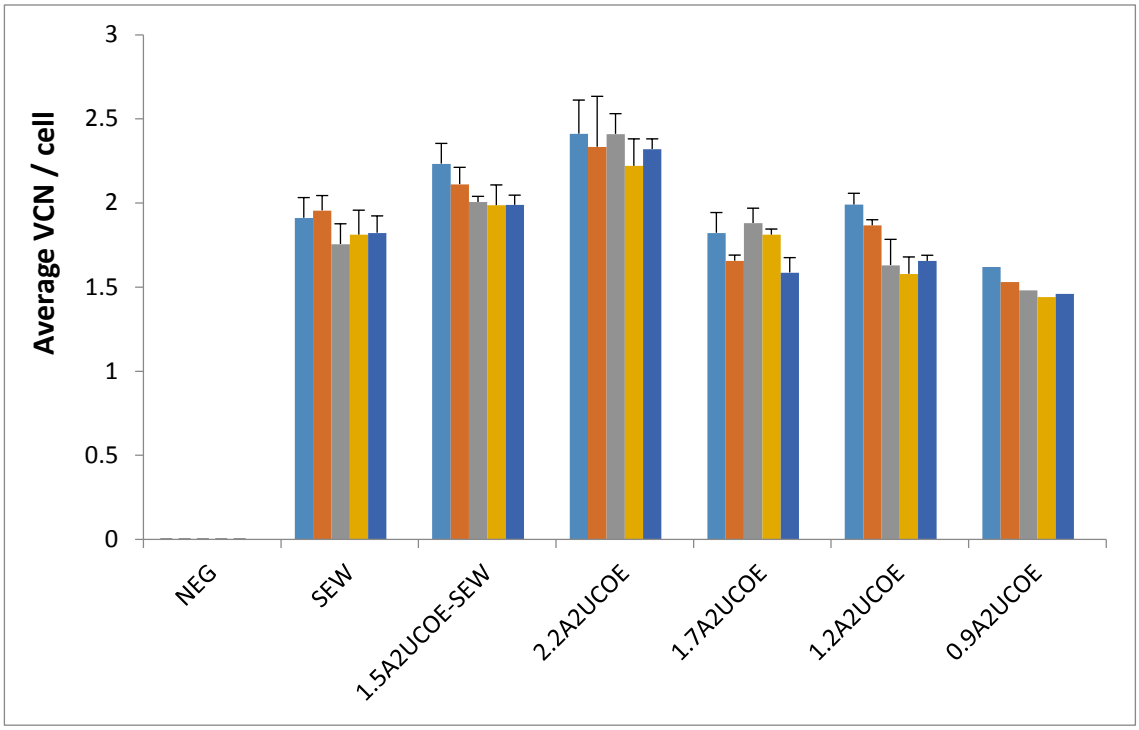


Figure 5.5: Novel candidate UCOEs offer only partial protection against silencing in differentiated F9 cells. F9 cells were transduced with novel candidate UCOE 1.2A2UCOE, 1.7A2UCOE and control SEW, 2.2kbA2UCOE and 1.5A2UCOE-SEW lentiviral vectors (Figure 5.1). Cells were analysed by flow cytometry to detect percentage of eGFP reporter gene expressing (eGFP+) cells, mean fluorescence intensity (MFI) and by RT-Q-PCR for average vector copy number (VCN) per cell. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 45 days post-transduction. (**A**) timecourse of percentage eGFP-positive cells; (Mean + SEM, n=4; **p<0.01). (**B**) As in (A) but showing MFI; (Mean + SEM, n=4; **p<0.01). (**C**) As in (A)/(B) but average VCN/cell; (Mean + SEM, n=4; **p<0.01).

5.6 Summary and Conclusions

The primary aim of the experiments conducted in this phase of the project was to determine whether the novel candidate sub-fragment UCOEs 1.2A2UCOE and 1.7A2UCOE (Figure 5.1) confirm with the dual divergent transcription model of UCOE function and are able to confer stability of expression on a linked heterologous promoter. Although the 1.5kb A2UCOE-SEW and 2.2kb UCOE can confer stability of expression on a linked ubiquitous (Williams, Mustoe et al. 2005, Zhang, Frost et al. 2010) or tissue-specific (Talbot, Waddington et al. 2009, Brendel, Müller-Kuller et al. 2011) promoter it does so in an orientation specific manner with the *CBX3* end of the element abutting the linked promoter. Our starting hypothesis was that this problem could be circumvented if both promoters of the selected gene pairs showed similar levels and variance of expression in a variety of tissues. That is the reason why uniformity of expression levels both across tissues but also in between the gene pair was taken into consideration. In this regard our results show that one of the sub-fragments of the candidate UCOEs 1.7A2UCOE is a good candidate to confer stability of expression (Figures 5.2 to 5.5).

In this phase of the project P19 and F9 cells were successfully transduced with two novel candidate UCOEs as well as three control LVs. Stability and levels of expression of these vectors were analyzed in both undifferentiated and neuroectodermal and endodermal differentiated P19 and F9 cells, respectively. Our study has extended previous investigations, which only looked at

expression of UCOE-based LVs profiles in undifferentiated P19 cells (Yoon, Lee et al. 2009, Zhang, Frost et al. 2010).The results provide comparison of the previously reported results of the 1.5kb A2UCOE vectors in P19 cells as well as that of the control vectors. The eGFP expression of the candidate UCOE vector 1.2A2UCOE decreased more rapidly than the 1.7A2UCOE. However, 1.5A2UCOE-SEW, 2.2kb A2UCOE and 1.7A2UCOE gave perfectly stable expression over the course of our experiment, as before (Zhang, Frost et al. 2010). Overall our results provide more evidence of the capacity of the A2UCOE, whether present in the configurations found in 1.5A2UCOE-SEW to provide stability and reproducibility of transgenes under the control of their innate *HNRPA2B1* promoter or driven by a linked heterologous promoter such as SFFV whose activity is protected from silencing and positional effect variegation by the dominant chromatin opening function of the UCOE.

Chapter 6
Results

Functional analysis of the A2UCOE with direct transgene expression from the *HNRPA2B1* promoter within murine ES cells

Chapter 6

Results: Functional analysis of the A2UCOE with direct transgene expression from the *HNRPA2B1* promoter within murine ES cells

6.1 Introduction and background

It has previously been shown that the 1.5kb core A2UCOE and sub-fragments thereof, can confer stability of expression on linked heterologous ubiquitous (Pfaff, Lachmann et al. 2013) (Ackermann, Lachmann et al. 2014) (Williams, Mustoe et al. 2005, Zhang, Frost et al. 2010) and tissue-specific (Müller-Kuller, Ackermann et al. 2015) promoters including within both embryonic and induced pluripotent stem cells and their differentiated progeny of all three germ layers. In Chapter 5 we showed that a 1.7kb sub-fragment of the fully functional 2.2kb A2UCOE retained the ability for stable expression in undifferentiated P19 and F9 cells as well as upon differentiation of these embryonal carcinoma cell lines down the neuroectoderm and parietal endoderm lineages, respectively. However, it has not as yet been established as to whether the 2.2A2UCOE and now also the 1.7A2UCOE vectors, which drive transgene expression directly off the innate *HNRPA2B1* promoter of this element, are also capable of stable expression in more physiological undifferentiated and differentiated murine embryonic stem (ES) cells. The aims of the experiments presented in this chapter were designed to test this possibility.

Note: the experiments presented in this chapter were conducted collaboratively with Prof Agi Grigoriadis and Ms Ewa Kania, Department of Craniofacial Development and Stem Cell Biology, King's College London.

6.2 LV used in this study

The LV constructs SEW (SFFV-eGFP-WPRE) and 1.5A2UCOE-SEW are as previously described in Chapter 3 (Figure 3.1). The 2.2A2UCOE, 1.7A2UCOE and 1.2A2UCOE vectors are as described in Chapter 5 (Figure 5.2). In the interests of clarity of presentation all these LVs are again illustrated here in Figure 6.1.

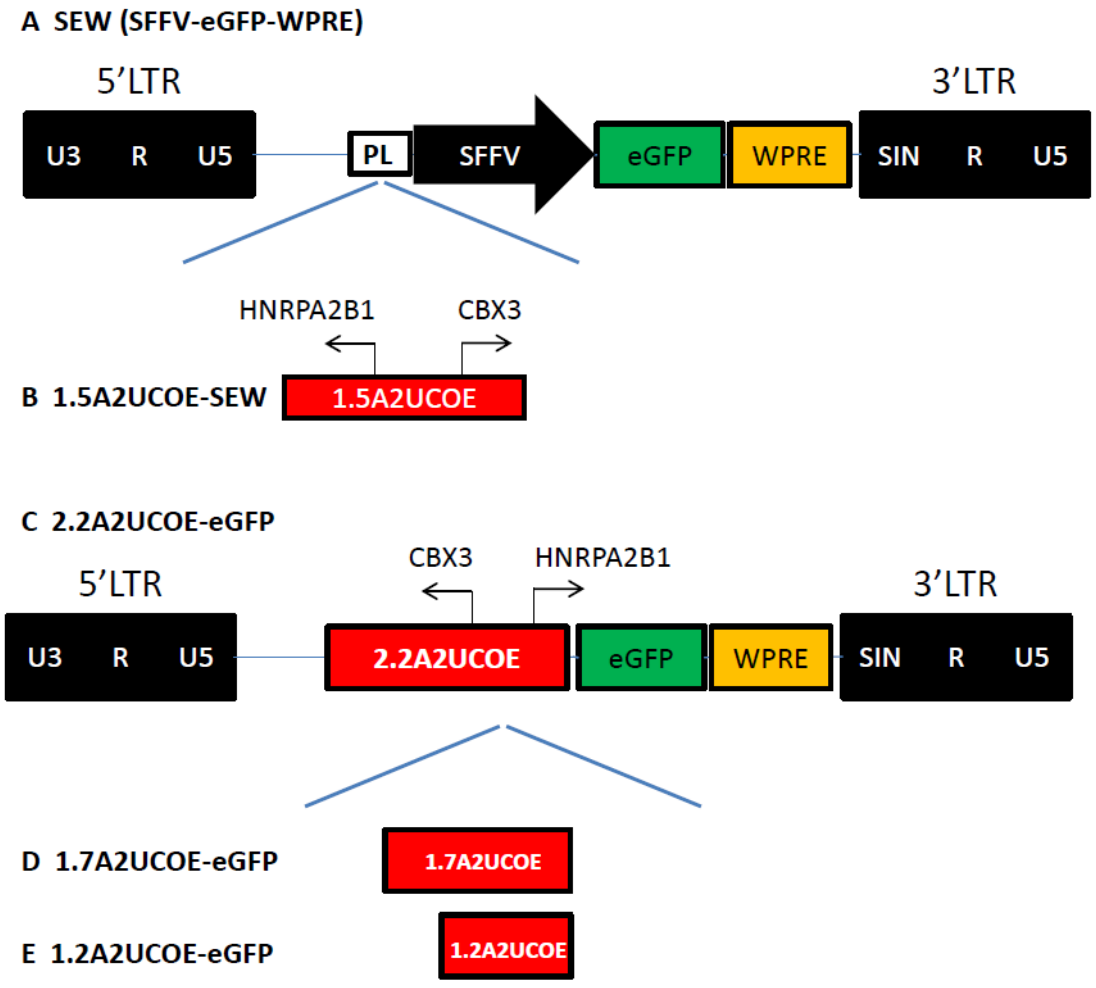

A SEW (SFFV-eGFP-WPRE)
5'LTR
3'LTR
U3
R
U5
PL
SFFV
eGFP
WPRE
SIN
R
U5
HNRPA2B1
CBX3
B 1.5A2UCOE-SEW
1.5A2UCOE
C 2.2A2UCOE-eGFP
CBX3
HNRPA2B1
5'LTR
3'LTR
U3
R
U5
2.2A2UCOE
eGFP
WPRE
SIN
R
U5
D 1.7A2UCOE-eGFP
1.7A2UCOE
E 1.2A2UCOE-eGFP
1.2A2UCOE

Figure 6.1. Illustration of the novel candidate UCOE and control lentiviral vectors. A. SEW: self-inactivating (SIN) LV containing a spleen focus forming virus (SFFV) promoter driving expression of an enhanced green fluorescent protein (eGFP) reporter gene with downstream woodchuck hepatitis virus post-transcriptional regulatory element (WPRE). **B.** 1.5A2UCOE-SEW: 1.5kb core *HNRPA2B1-CBX3* UCOE (1.5A2UCOE) was inserted into a polylinker (PL) cloning site upstream of the SFFV promoter within SEW. **C.** 2.2A2UCOE-eGFP: 2.2kb *HNRPA2B1-CBX3* UCOE (2.2A2UCOE) directly expressing a GFP reporter gene off the innate *HNRPA2B1* promoter. **D, E.** 1.7A2UCOE-eGFP and 1.2A2UCOE-eGFP: 5' (*CBX3* end) deletion series of the 2.2A2UCOE consisting of 1.7kb and 1.2kb sub-fragments. LTR: long terminal repeat. Horizontal arrows denote direction of transcription.

6.3 Lentiviral vector titration in HEK293T cells

Cells were collected for analysis by flow cytometry three days post-transduction of HEK293T cells with serial dilutions of viral stocks collected at both 2 and 3 days post-transfection from each production run for all LVs (Table 6.1).

LV	Volume (µL) of viral vector stock			
	2 µL	0.2 µL	0.02 µL	0.002 µL
	Percentage eGFP+ cells 1st / 2nd viral harvest			
SEW	84.15 / 62.45	15.12 / 11.24	4.2 / 1.66	0.92 / 0.44
1.5A2UCOE-SEW	72.22 / 61.35	13.11 / 10.22	3.7 / 1.55	0.71 / 0.35
2.2A2UCOE	75.3 / 62.44	15.22 / 9.58	3.66 / 1.25	0.83 / 0.51
1.7A2UCOE	68.55 / 61.21	14.12 / 8.45	3.01 / 1.11	0.62 / 0.14
1.2A2UCOE	73.48 / 65.33	14.91 / 11.2	3.25 / 1.26	0.78 / 0.18

Table 6.1. Titration of lentiviral vector (LV) preparations. A $2x10^5$ aliquot of HEK293T cells were transduced with 0.002-2 µL of a given LV stock from the first and second harvest of virus from culture supernatant during the time of preparation. Analysis of eGFP-positive cells at 3-days post-transduction was by flow cytometry. LV constructs are as described in Figure 6.1.

6.3.1 Calculation of LV titre

Viral titre (Table 6.2) was calculated based on the dilution of each harvest that gave a percentage of eGFP-positive cells between 1% and 10%. Pools of cells with percentages of eGFP-positive cells higher than 10% are very likely to contain multiple integrations of the vector, while very low scores could be false positives.

On day one, $2x10^5$ 293T cells were seeded in each well. Thus, the percentage of eGFP-positive cells from flow cytometry analysis, reflects the percentage of the initial cell population that was successfully transduced and hence the number of infectious units added in the specified well. The dilution factor or the volume of the lentivirus preparation that was used is known, so the calculation of the number of infectious units per ml from that point onwards is straightforward.

LV	Titre (iu/mL)	
	1st harvest	2nd harvest
SEW	2.54 x 10^8	1.81 x 10^8
1.5A2UCOE-SEW	2.22 x 10^8	1.46 x 10^8
2.2A2UCOE	2.73 x 10^8	1.36 x 10^8
1.7A2UCOE	2.47 x 10^8	1.16x 10^8
1.2A2UCOE	2.11 x 10^8	1.05 x 10^8

Table 6.2. Lentiviral vector (LV) titres. Viral titre as infectious units per ml (iu/mL) was calculated based on the dilution of 1st and 2nd harvest of virus that gave a percentage of eGFP-positive HEK293T cells between 1% and 10% (Table 6.1). LV constructs are as described in Figure 6.1.

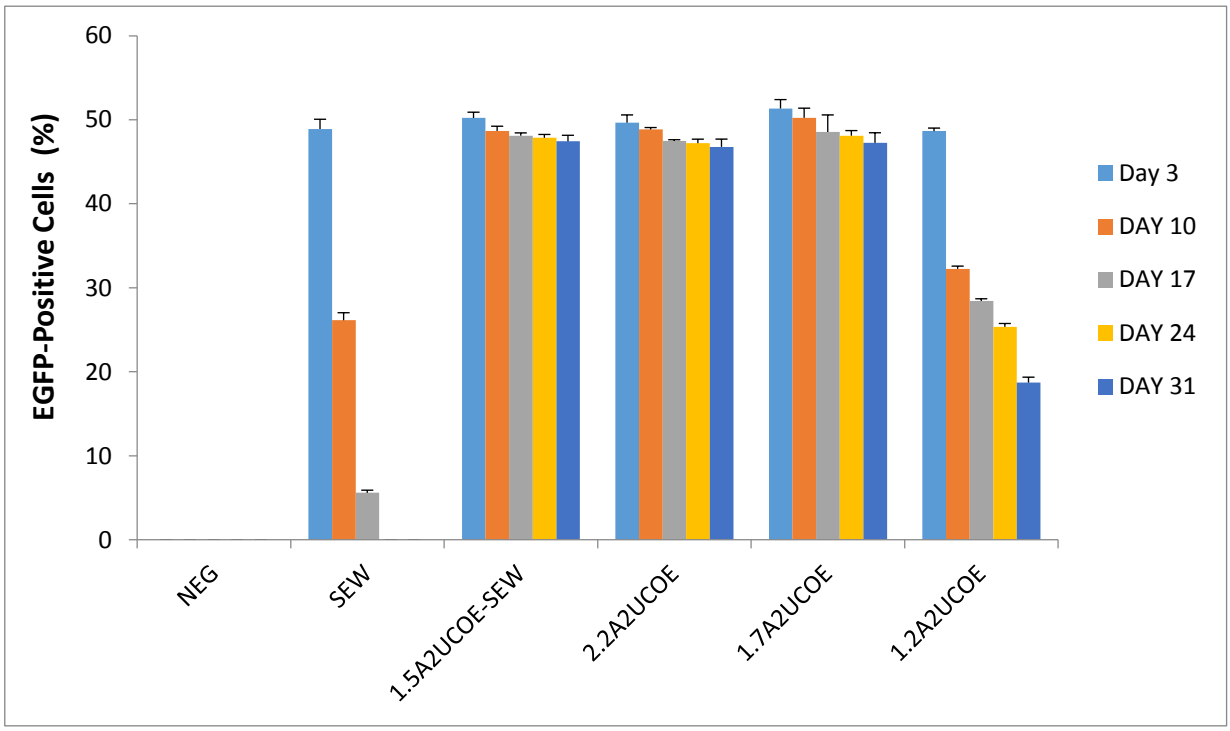
6.4 Functional assay of candidate 1.2A2UCOE and 1.7A2UCOE LVs in undifferentiated murine ES cells

Murine ES cells were transduced with the generated LVs at an MOI of 3, with the intention to start the experiment within the range of 40-60% eGFP-positive cells in all pools. The transduced cells and untransduced negative controls were then propagated and assayed as a time course by flow cytometry to determine the percentage of eGFP-positive cells and mean fluorescence intensity (MFI) for each repeat transduction for each vector. Cell cultures were analysed every 7 days from day 3 post-transduction and extending to 31 days. In addition, DNA was extracted from cells at each time point for qPCR analysis in order to determine the average VCN per cell.

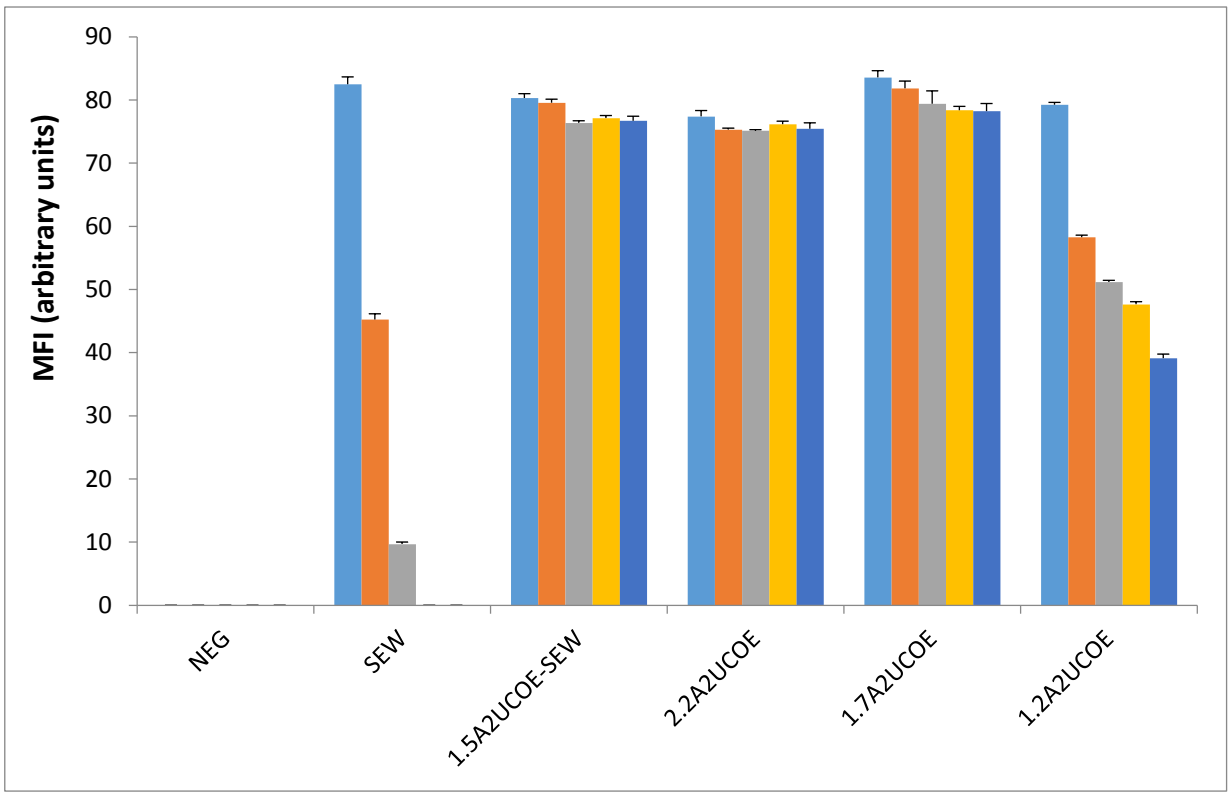
Figure 6.2A shows the flow cytometry time course results depicting percentage of eGFP+ cells. Initial transduction efficiency was between 45-52% eGFP+ cells in all cases. However, expression from SEW rapidly declined from 48% to 5% eGFP+ cells within 17 days. In contrast, the proportion of eGFP+ cells from the 1.5A2UCOE-SEW and 2.2A2UCOE vectors remained completely stable over the 31-day period of culture. The 1.7A2UCOE LV displayed the same stable pattern of expression to that seen with the 2.2A2UCOE vector over the time course of the experiment. By comparison, the percentage of eGFP+ cells transduced with the 1.2A2UCOE construct decreased over time although much more gradually than that seen with SEW. The expression of eGFP in SEW transduced cells dropped by almost 80% after 2 weeks post-transduction whereas the percentage of eGFP-positive cells decreased by only 30% to 40% in the case of the 1.2A2UCOE vector over the same time period. The values of

mean fluorescent intensity (MFI) for all vectors paralleled the eGFP expression results (Figure 6.2B). MFI was stable in the 1.5A2UCOE-SEW, 2.2A2UCOE and 1.7A2UCOE transduced cells and unstable in the case of SEW and 1.2A2UCOE. Finally, average VCN per cell was unaltered during the course of the experiment clearly demonstrating that loss of expression from the SEW and 1.2A2UCOE constructs was due to transgene silencing and not vector loss.

A)


EGFP-Positive Cells (%)
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
1.7A2UCOE
1.2A2UCOE
Day 3
DAY 10
DAY 17
DAY 24
DAY 31

B)


MFI (arbitrary units)
90
80
70
60
50
40
30
20
10
0
NEG
SEW
1.5A2UCOE-SEW
2.2A2UCOE
1.7A2UCOE
1.2A2UCOE

C)

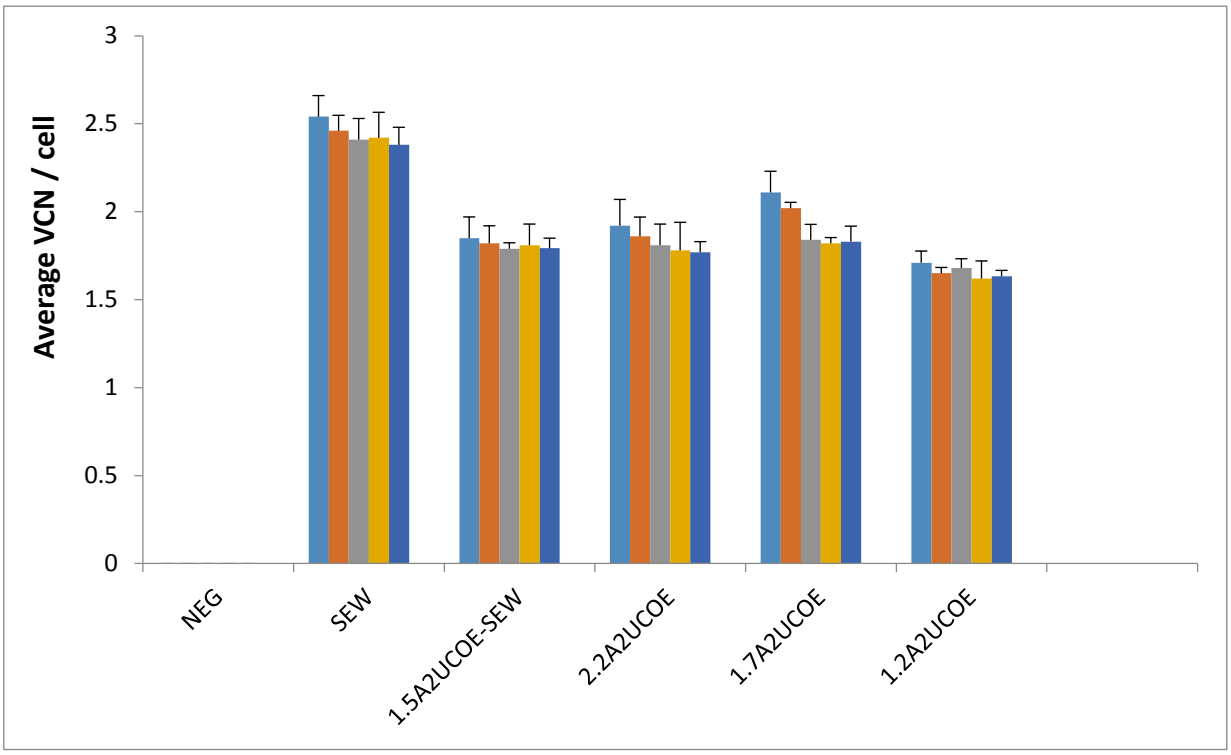
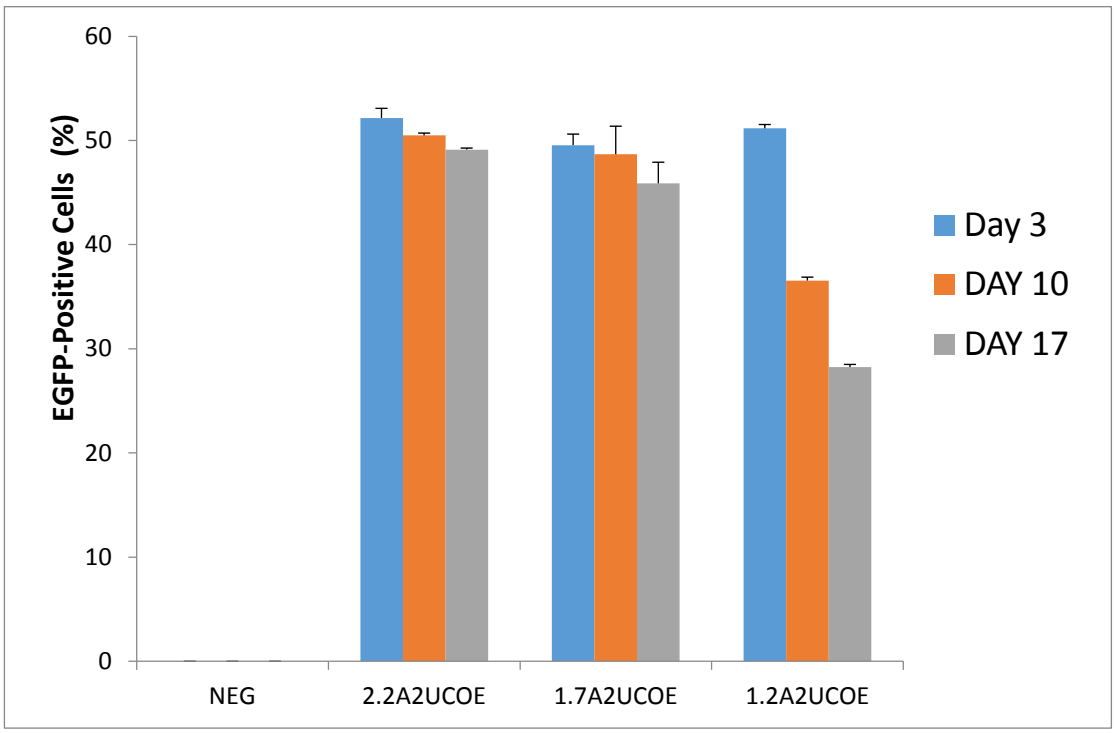


Figure 6.2. The 2.2A2UCOE and 1.7A2UCOE offer complete protection against silencing in undifferentiated murine ES cells. Murine ES cells were transduced with the 2.2A2UCOE, 1.7A2UCOE, 1.2A2UCOE and control SEW and 1.5A2UCOE-SEW LVs (Figure 6.1). Cells were analysed by flow cytometry to detect percentage of eGFP reporter gene expressing (eGFP+) cells, mean fluorescence intensity (MFI) and by qPCR for average vector copy number (VCN) per cell. Data shows combined results from three independent transductions for each vector, plus negative control (NEG), over a period of 3 to 31 days post-transduction. (**A**) timecourse of percentage eGFP-positive cells; (Mean + SEM, n=4; **p<0.01). (**B**) As in (A) but showing MFI; (Mean + SEM, n=4; **p<0.01). (**C**) As in (A)/(B) but average VCN/cell; (Mean + SEM, n=4; **p<0.01).

6.5 Functional analysis of 2.2A2UCOE, 1.7A2UCOE and 1.2A2UCOE vectors in differentiated murine ES cells

The 2.2A2UCOE and 1.7A2UCOE vectors displayed complete stability of expression in undifferentiated ES cells (Figure 6.2). We next sought to determine if these LVs retained their capability for stable expression upon differentiation of these cells. ES cells were transduced with the 2.2A2UCOE, 1.7A2UCOE and 1.2A2UCOE series of vectors and cultured under non-adhesive conditions on uncoated bacterial Petri dishes to induce embryoid body formation, which would contain cell types representative of all three germ layers. Analysis by flow cytometry to determine percentage eGFP+ cells and MFI was conducted on the cell cultures prior to induced embryoid body differentiation at 3-days post-transduction and at day 10 and 17 following differentiation. The results (Figure 6.3) show that the 2.2A2UCOE and 1.7A2UCOE vectors essentially retain the same level of eGFP+ cells (Figure 6.3A) and MFI (Figure 6.3B) up to the 17 day period of the experiment. In contrast, expression from the 1.2A2UCOE LV declined over time, reducing to approximately 40% of eGFP+ starting values by day 17 with a parallel drop in MFI. Thus, by these preliminary results the expression profile seen with the 2.2A2UCOE, 1.7A2UCOE and 1.2A2UCOE LVs in embryoid body differentiated ES cells was similar to that seen in undifferentiated cells (Figure 6.2); that is, the 2.2A2UCOE and 1.7A2UCOE vectors retained stability of expression whilst the function of the 1.2A2UCOE was unstable following ES cell differentiation to embryoid bodies.

A)


EGFP-Positive Cells (%)
60
50
40
30
20
10
0
NEG
2.2A2UCOE
1.7A2UCOE
1.2A2UCOE
Day 3
DAY 10
DAY 17

B)

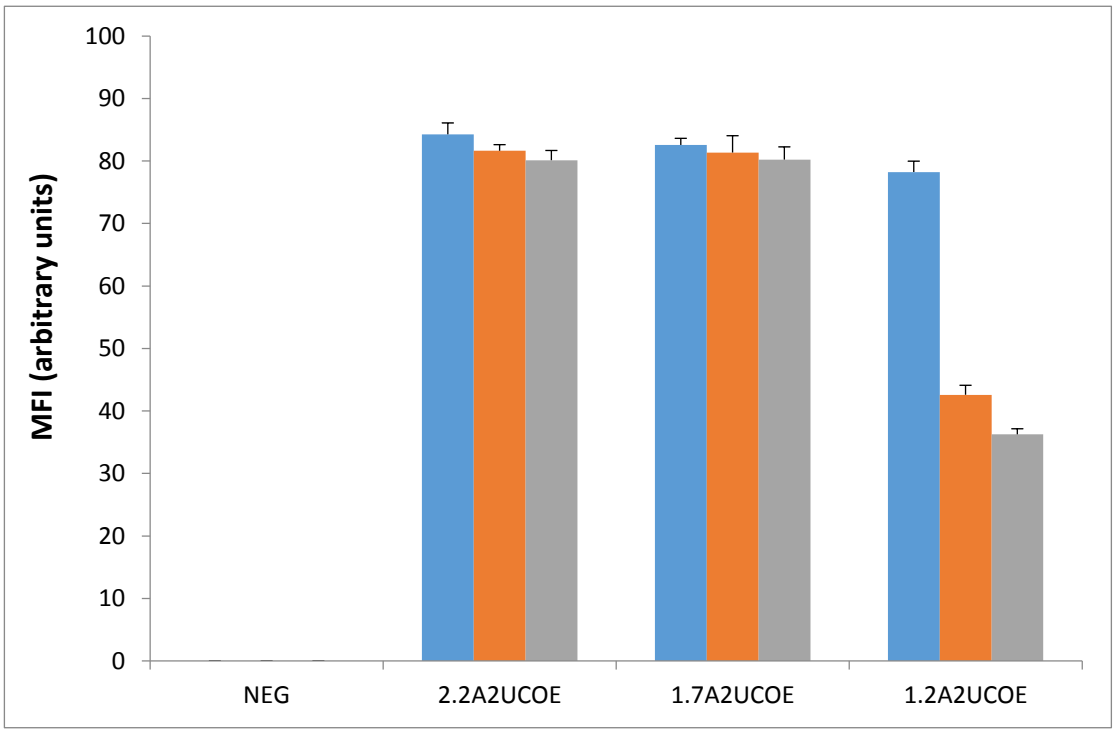

MFI (arbitrary units)
100
90
80
70
60
50
40
30
20
10
0
NEG
2.2A2UCOE
1.7A2UCOE
1.2A2UCOE

Figure 6.3. The 2.2A2UCOE and 1.7A2UCOE vectors but not the 1.2A2UCOE construct retain stability of expression following murine ES cell differentiation to embryoid bodies. Murine ES cells were transduced with the 2.2A2UCOE, 1.7A2UCOE and 1.2A2UCOE LVs (Figure 6.1) and at 3-days post-transduction transferred to differentiation medium in non-adherent uncoated bacterial Petri dishes to induce embryoid body formation. Cells were analysed prior to (day 3) and after (days 10 and 17) induced embryoid body differentiation by flow cytometry to measure percentage of eGFP reporter gene expressing (eGFP+) cells and mean fluorescence intensity (MFI). Data shows combined results from three independent transductions for each vector, plus negative (untransduced) control (NEG). (**A**) timecourse of percentage eGFP-positive cells; (Mean + SEM, n=4; **p<0.01). (**B**) As in (A) but showing MFI; (Mean + SEM, n=4; **p<0.01).

6.6 Summary and Conclusions

The aim of this phase of the project was to evaluate whether A2UCOE-based LVs where the gene of interest is transcribed directly off the innate *HNRPA2B1* promoter, retain stability of expression in undifferentiated and embryoid body differentiated murine ES cells. To this end ES cells were transduced and analysed with the 2.2A2UCOE, 1.7A2UCOE and 1.2A2UCOE vectors (Figure 6.1).

Our results clearly show that as in the case for these vectors in murine P19 and F9 embryonal carcinoma cells (see Chapter 3), the 2.2A2UCOE and 1.7A2UCOE constructs retained a full capability for stable expression in undifferentiated ES cells (Figure 6.2). In contrast the 1.2A2UCOE vector retained at best only partial stability of expression although at a significantly superior level to that of the SFFV-based SEW silencing control LV. Preliminary results in ES cells induced to undergo differentiation to embryoid bodies, which contain cells representative of all three germ layers, also suggest that 2.2A2UCOE and 1.7A2UCOE but not 1.2A2UCOE retain their capability for stable expression (Figure 6.3).

These data represent the first evidence that A2UCOE-based LV constructs with expression off the innate *HNRPA2B1* promoter provide reproducible and stable expression within undifferentiated and differentiated ES cells. Thus, vectors employing the 1.7A2UCOE-driven transgenes can be considered as very good

candidates for research and therapeutic purposes where reproducible and long-term stability of expression is required when targeting ES cells and their differentiated progeny.

Chapter 7

Discussion

Chapter 7

Discussion

The integrating gammaretroviral and lentiviral class of vectors remain the best option for achieving stable retention and expression of a therapeutic gene. This is especially the case when targeting mitotic stem cell populations. Indeed, over the 15 years we have seen successful outcomes in clinical trials using these classes of vectors targeting HSCs via an *ex vivo* approach. Severe combined immunodeficiency disease (SCID), particularly SCID-X1, SCID-ADA (Cavazzana-Calvo, Fischer et al. 2012, Cavazza, Moiani et al. 2013) and Wiskott-Aldrich syndrome (WAS) and chronic granulomatous disease (CGD) ((Ott, Seger et al. 2007, Aiuti, Bacchetta et al. 2012) have been successfully treated by genetic augmentation of the patient's HSC with gammaretroviral vectors. More recently, LVs have been successfully employed to treat the inherited demyelinating conditions X-(Cartier, Hacein-Bey-Abina et al. 2009) and MLD (Biffi, Montini et al. 2013) as well as WAS (Bosticardo, Ferrua et al. 2014).

However, one is faced with two problems that always need to be addressed when using these integrating classes of viral vectors. First, insertional mutagenesis and second epigenetic-mediated therapeutic transgene silencing (Antoniou, Skipper et al. 2013). The problem of insertional mutagenesis by gammaretroviral integration causing inadvertent upregulation of host cell proto-oncogenes and oncogenesis, was observed in 5 out of a total of 20 patients treated for SCID-X1 (Cavazzana-Calvo, Fischer et al. 2012, Cavazza, Moiani et al. 2013). Transgene silencing by promoter

DNA methylation ultimately led to failure of initially successful gene therapy in two CGD patients (Ott, Seger et al. 2007, Aiuti, Bacchetta et al. 2012).

The overall theme of this thesis is concerned with addressing the problem of silencing of LV-delivered therapeutic transgenes by the use of ubiquitous chromatin opening elements (UCOEs). The UCOE from the human *HNRPA2B1-CBX3* housekeeping gene locus (A2UCOE) has now seen extensive use within LVs, which has demonstrated its ability to provide unprecedented reproducible and stable transgene expression *in vivo* especially when targeting stem cell populations (Antoniou, Skipper et al. 2013).

The proposed mechanism by which the A2UCOE is able to confer both reproducibility and stability of expression independent of the site of transgene integration, consists of two components; a region of extended CpG methylation-free region coupled with divergent transcription with an inherent chromatin opening capability from the *HNRPA2B1* and *CBX3* promoters (Allen and Antoniou 2007). The aim of the work conducted under this thesis was to functionally investigate these proposed mechanisms of action. Ultimately it was hoped to identify a minimal, fully functional A2UCOE for incorporation within LVs thereby maximising space for the therapeutic gene.

7.1 Novel candidate *SETD3-CCNK* and artificial *RPS11-HNRPA2B1* dual divergent transcriptional elements possess only a partial UCOE function

It has previously been reported that the placement of the core 1.5kb or 1.2kb A2UCOE upstream of heterologous promoters would confer stability of expression in an orientation specific manner; that is, when the *CBX3* end of the A2UCOE was placed next to the heterologous ubiquitous SFFV (Zhang, Frost et al. 2010) and EF1α (Pfaff, Lachmann et al. 2013) promoters or tissue-specific MRP8 (Brendel, Müller-Kuller et al. 2012) promoter. The reason put forward to try and explain this orientation-bias in A2UCOE function was because the relatively weak transcriptional activity of the *CBX3* promoter compared to that of *HNRPA2B1*. Thus, divergent transcription from the stronger *HNRPA2B1* promoter would provide a better barrier against repressive epigenetic (DNA methylation, histone modification) processes spreading into the transgene region and thus silencing expression. Our first series of experiments were designed to test this hypothesis by seeking to identify and construct UCOEs with a dual divergent transcriptional configuration but which may function equally in either orientation when linked to heterologous promoters. Our starting hypothesis was that any problem in orientation bias in UCOE-heterologous promoter combination function could be circumvented if both promoters of the selected gene pairs showed similar levels and variance of expression between tissues.

The first novel candidate UCOE selected was the *SETD3-CCNK* housekeeping gene pair as these two dual divergently transcribed pair of genes are both relatively uniformly expressed across 40 different human tissues and cell types (She, Rohl et al. 2009). In order to directly test the hypothesis that the orientation bias of the A2UCOE

when linked to heterologous promoters was due to the weak transcriptional activity of *CBX3*, an artificial UCOE was constructed whereby the *CBX3* component of the A2UCOE was replaced with the single promoter-CpG island of the *RPS11* housekeeping gene but still in a divergently transcribing configuration. The *RPS11* element was again chosen as this gene is expressed at a uniformly high level across different human tissues (She, Rohl et al. 2009).

These novel candidate UCOE *SETD3-CCNK* and *RPS11-HNRPA2B1* elements were linked upstream of the highly silencing-prone SFFV promoter (Figure 3.1) to evaluate their ability to prevent repression of this element in undifferentiated and differentiated P19 and F9 cells. In one respect the results we obtained are encouraging in that both orientations of the *SETD3-CCNK* element (SET-CCN-SEW, CCN-SET-SEW) and the artificial *HNRPA2B1-RPS11* combination (B1-RPS-SEW, RPS-B1-SEW) showed similar abilities to confer stability of expression from the SFFV promoter independent of orientation in both cell lines and all conditions (Figures 3.4, 3.6, 3.8, 3.10). However, our results also show that neither the *SETD3-CCNK* nor the *HNRPA2B1-RPS11* elements were as effective as the prototypical 1.5A2UCOE at negating silencing of the linked SFFV promoter. The 1.5A2UCOE-SEW construct gave perfectly stable expression over the course of our experiment in not only undifferentiated P19 cells as previously reported (Zhang, Frost et al. 2010), but also in F9 cells. Furthermore, the 1.5A2UCOE-SEW vector maintained stable expression upon neuroectodermal and endodermal differentiation of P19 and F9 cells, respectively. Overall, our results from these experiments provided further evidence of the potent capacity of the A2UCOE present in the configuration found in

1.5A2UCOE-SEW LV to provide stability and reproducibility of transgene expression when linked to a heterologous promoter such as SFFV.

7.2 Sub-regions of the A2UCOE fail to retain function

A number of publications have questioned the need for divergent or even any associated promoter activity for UCOE function. Bandarayake and colleagues claimed that a 700bp fragment from the 3' end of intron I of *CBX3* (Figure 4.1, 0.7UCOE), and thus devoid of promoter activity, was able to confer stability of expression on a linked CMV-GFP reporter construct from within an LV context in CHO cells albeit at high (8-10) average vector copies per cell (Bandaranayake, Correnti et al. 2011). We thus sought to test this same 0.7UCOE "Daedalus" element within our standard assay system by placement upstream of the SFFV promoter in forward and reverse orientations within the SEW LV (Figure 4.3; Daedalus-F, Daedalus-R) and functional analysis in P19 and F9 cells. Our results (Figures 4.4-4.7) clearly show that this Daedalus 0.7UCOE was unable to confer stability of expression on the linked SFFV promoter in either orientation in both undifferentiated and differentiated P19 and F9 cells. It was clear from our results that the reason for apparent stability of expression seen in CHO cells was due to the high average vector copy number per cell and that most integration events were in fact being silenced as evidenced by the dramatic drop in MFI at 14 days post-transduction (Bandaranayake, Correnti et al. 2011). These data demonstrate the importance of evaluating stability of expression from LVs at low vector copy per cell otherwise misleading results may be obtained, especially when considering applications within a gene therapy context.

A report by Thomson and colleagues (Thomson, Skene et al. 2010) suggested that transgenes consisting of a CpG-rich DNA fragment but lacking a promoter are sufficient to establish a methylation-free region with associated active histone modification marks in the absence of RNA polymerase II. Based on these findings, we sought to analyse the most CpG dinucleotide dense sub-regions of the A2UCOE, especially those lacking either the *HNRPA2B1* or *CBX3* promoters, for UCOE function. The most CpG-dense region of the A2UCOE CpG island lies within the first intron of *CBX3* immediately downstream of the two alternative first exons (Figure 4.1; 945UCOE, 527UCOE, 455UCOE) and was thus chosen for functional analysis. This 945bp region was divided into two sub-fragments of 455bp and 527bp and all three linked individually upstream of the SFFV promoter within the SFFV-eGFP (SEW) vector (Figure 4.3; 945UCOE, 527UCOE, 455UCOE). Our results show that despite the very high CpG density of these fragments, none of them were able to confer stability of expression on the linked SFFV promoter in both undifferentiated and differentiated P19 and F9 cells (Figures 4.14-4.19). Thus, CpG density alone is clearly not sufficient to establish a methylation-free open chromatin structure and that some other feature or features are also required. The missing component can of course be the lack of associated promoter and transcriptional activity with the 945UCOE, 527UCOE and 455UCOE elements. Alternatively, the length of the CpG island region may have been reduced below a critical length required for it to maintain its methylation-free, open chromatin and transcriptionally permissive structure (Allen and Antoniou 2007).

It has also been reported that sub-fragments of CpG islands associated with the promoters of developmentally regulated genes retain their ability to maintain

appropriate DNA methylation status in a transgene context (Lienert, Wirbelauer et al. 2011). This included some CpG island sub-fragments that lacked their cognate promoter region and gave rise to them being designated as "methylation-determining regions" (MDRs). Based on these interesting observations we decided to ascertain if a similar MDR was present in the housekeeping gene region of the A2UCOE. A candidate MDR within the A2UCOE region was initially sought by a bioinformatics investigation for key transcription factor (Sp1, CTCF, USF) binding sites that are thought to constitute an MDR signature (Lienert, Wirbelauer et al. 2011). This revealed a 0.9kb core fragment extending from the first exons of *CBX3* and *HNRPA2B1* and thus including both promoter regions as being a possible MDR (data not shown). As in previous experiments, we linked the 0.9kb core A2UCOE region encompassing the hypothetical MDR (Figure 4.1; 0.9UCOE) upstream of the SFFV promoter within the SEW LV (Figure 4.3; 0.9UCOE-F and 0.9UCOE-R) to test its ability to confer stability of expression in our proven differentiated and undifferentiated P19 and F9 cell assay. Our results show that the MDR candidate 0.9kb A2UCOE region is able to confer significant but still partial protection against silencing of the linked SFFV promoter when compared to the larger 1.5A2UCOE core fragment in both undifferentiated and differentiated P19 (Figures 4.8, 4.10, 4.12) and F9 (Figures 4.9, 4.11, 4.13) cells. Thus, although the 0.9kb core A2UCOE promoter fragment may constitute an MDR, it does not retain a full UCOE capability as does the larger 1.5kb A2UCOE, which extends further at both the *HNRPA2B1* and *CBX3* ends. As in the case of the 945UCOE, 527UCOE and 455UCOE fragments, this difference in UCOE activity between these two A2UCOE core elements may be due the reduction in the length of the CpG island beyond a crucial point within the 0.9kb fragment such that it is unable to stably maintain its normally inherent open chromatin

structure (Allen and Antoniou 2007). Alternatively, binding sites for additional transcription factors that are crucial for UCOE function may have been deleted in reducing A2UCOE length from 1.5kb to 0.9kb.

In conclusion, none of the novel candidate UCOE sub-fragments of the A2UCOE region either with (0.9UCOE-F, 0.9UCOE-R) or without (Daedalus-F, Daedalus-R; 455UCOE, 527UCOE, 945UCOE) associated promoter activity retained a full UCOE capability. Only the 0.9UCOE-F and especially 0.9UCOE-R constructs that encompass the promoters and transcriptional start sites of *HNRPA2B1* and *CBX3* possessed a significant but still partial ability to negate transgene silencing. Thus, despite published evidence to the contrary we were unable by these targeted experiments to identify an A2UCOE sub-region devoid of promoter activity that still retained a full UCOE function.

7.3 A 1.7kb sub-fragment of the 2.2kb A2UCOE retains full activity

Prior to the commencement of this project it had been shown that the innate *HNRPA2B1* promoter of the 2.2kb A2UCOE (2.2A2UCOE; Figure 5.1) can provide stable transgene expression from both plasmid (Antoniou, Harland et al. 2003) and lentiviral (Zhang, Thornhill et al. 2007) vectors. (See also, for example, Figures 4.8 and 4.9). In contrast, a 0.9kb A2UCOE (0.9A2UCOE; Figure 5.1) again where transcription was directly from the *HNRPA2B1*promoter, did not provide stable expression in P19 cells transduced with a LV containing a 0.9A2UCOE-eGFP cassette (Knight, Zhang et al. 2012). We, therefore, investigated as to whether lengths

of the A2UCOE between the 2.2A2UCOE and 0.9A2UCOE constructs retained full activity. This entailed building deletion constructs of 1.7kb and 1.2kb of the 2.2A2UCOE (Figure 5.1) and linkage to an eGFP reporter gene with expression from the *HNRPA2B1* promoter (Figure 5.2). These vectors were tested for UCOE function in our standard murine embryonal carcinoma P19 and F9 cell assay system with stability, potency, and reproducibility of expression of the eGFP transgene were analyzed both before and after differentiation.

Our results clearly show that the 1.7A2UCOE but not the 1.2A2UCOE construct provides comparable stability of expression to that of the full 2.2A2UCOE in both undifferentiated and differentiated P19 (Figures 5.3, 5.5, 5.7) and F9 (Figures 5.4, 5.6, 5.8) cells. We also showed that the smaller 0.9A2UCOE fragment previously shown not to retain stability of expression in undifferentiated P19 cells (Knight, Zhang et al. 2012), is also incapable of stable expression in undifferentiated F9 cells and in both differentiated P19 and F9 cells.

The inability of the 1.2A2UCOE and 0.9A2UCOE vectors to maintain stability of expression is despite that fact that both constructs retain a dual *CBX3-HNRPA2B1* divergently transcribed promoter configuration (Figure 5.1). In addition, our finding of instability of expression from the 1.2A2UCOE appears at odds with those of Müller-Kuller and colleagues (Müller-Kuller, Ackermann et al. 2015). They reported that a 0.7kb sub-fragment of the A2UCOE encompassing just the *CBX3* promoter and alternative first exons (designated as "CBX3-UCOE"), is capable of driving stable albeit very low levels of transcription and also able to stabilise expression from linked

heterologous SFFV and MRP8 promoters in P19, ES, HSC cells and derivatives. However, as the 0.7kb CBX3-UCOE constitutes the *CBX3* half and, therefore, a continuous part of the 1.2A2UCOE, it would be expected that expression from the naturally associated *HNRPA2B1* promoter would also be stabilised, which is clearly not the case in P19 (Figures 5.3, 5.5, 5.7), F9 (Figures 5.4, 5.6, 5.8) and ES (Figures 6.2 and 6.3) cells. The difference, of course, between our 1.2A2UCOE and the CBX3-UCOE-SFFV/MRP8 constructs tested by Müller-Kuller and colleagues, is that the direction of transcription of *CBX3* in 1.2A2UCOE is divergent to *HNRPA2B1* (Figure 6.1) whereas it is in the same direction in the case of both CBX3-UCOE-SFFV and CBX3-UCOE-MRP8 (Müller-Kuller, Ackermann et al. 2015). At face value these observations clearly question the requirement for divergent transcription for A2UCOE function although further experiments involving, for example reversing the orientation of the 0.7kb CBX3-UCOE region within the 1.2A2UCOE construct, are required to address the molecular basis of this discrepancy.

What our results with the 1.7A2UCOE and 1.2A2UCOE constructs suggest is that elements and/or functions in addition to divergent transcription are required for full A2UCOE function and that these are lost upon deletion of the 500bp *CBX3* region between the 1.7kb and 1.2kb endpoints of the 1.7A2UCOE and 1.2A2UCOE constructs. What these elements and processes might be is currently unknown but could include either deletion of crucial transcription factor binding sites or reduction of the methylation-free CpG island below a critical length, or both.

7.4 The 2.2A2UCOE and 1.7A2UCOE vectors retain full UCOE activity in murine embryonic stem cells

Given that the 1.7kb sub-fragment of the fully functional 2.2kb A2UCOE retained the ability for stable expression in undifferentiated and differentiated P19 and F9 cells (see Chapter 5), we also evaluated as to whether the 2.2A2UCOE and 1.7A2UCOE vectors are also capable of stable expression in more physiological undifferentiated and differentiated murine ES cells. Our results clearly show that both the 2.2A2UCOE and 1.7A2UCOE constructs retained a full capability for stable expression in undifferentiated and embryoid body differentiated ES cells (Figures 6.2 and 6.3). In contrast the 1.2A2UCOE vector retained at best only partial stability of expression although at a significantly superior level to that of the SFFV-based SEW silencing control LV (Figures 6.2 and 6.3). These data represent the first evidence that A2UCOE-based LV constructs with expression off the innate *HNRPA2B1* promoter provide reproducible and stable expression within undifferentiated and differentiated ES cells.

In summary, attempts in this project to identify and construct novel dual divergently transcribed promoter elements to match the UCOE capability of A2UCOE were only partially successful. In addition, a functional dissection of the A2UCOE by deletion analysis of the fully functional element failed to identity any sub-fragments that were CpG rich but devoid of one or both of the *HNRPA2B1* and *CBX3* promoters, which retained UCOE function. Only a 0.9kb core fragment spanning both *CBX3* and *HNRPA2B1* promoters was found to be significantly but partially protective against transcriptional silencing. Finally, we successfully reduced the size of fully functional

A2UCOE designed for expression off the innate HNRPA2B1 promoter to 1.7kb, reduction of 500bp over the previous best of 2.2kb.

Overall, our results do not dispute the two component model of A2UCOE function (dual divergent transcription from within an extended CpG island). In addition, from a practical perspective our finding that the 1.7A2UCOE retains the same stability of expression as the larger 2.2A2UCOE element suggests that it can effectively replace the latter within therapeutic LV constructs allowing a greater capacity for the gene of interest.

Chapter 8

References

References

Ackermann, M., N. Lachmann, S. Hartung, R. Eggenschwiler, N. Pfaff, C. Happle, A. Mucci, G. Göhring, H. Niemann and G. Hansen (2014). "Promoter and lineage independent anti-silencing activity of the A2 ubiquitous chromatin opening element for optimized human pluripotent stem cell-based gene therapy." Biomaterials **35**(5): 1531-1542.

Aiuti, A., R. Bacchetta, R. Seger, A. Villa and M. Cavazzana-Calvo (2012). "Gene therapy for primary immunodeficiencies: Part 2." Current Opinion in Immunology **24**(5): 585-591.

Alexander, S. L., W. T. Linde-Zwirble, W. Werther, E. E. Depperschmidt, L. J. Wilson, R. Palanki, N. Saroj, S. L. Butterworth and T. Ianchulev (2007). "Annual rates of arterial thromboembolic events in medicare neovascular age-related macular degeneration patients." Ophthalmology **114**(12): 2174-2178.

Allen, M. L. and M. Antoniou (2007). "Correlation of DNA methylation with histone modifications across the HNRPA2B1-CBX3 ubiquitously-acting chromatin open element (UCOE)." Epigenetics **2**(4): 227-236.

Antoniou, M., L. Harland, T. Mustoe, S. Williams, J. Holdstock, E. Yague, T. Mulcahy, M. Griffiths, S. Edwards and P. A. Ioannou (2003). "Transgenes encompassing dual-promoter CpG islands from the human TBP and HNRPA2B1 loci are resistant to heterochromatin-mediated silencing." Genomics **82**(3): 269-279.

Antoniou, M., L. Harland, T. Mustoe, S. Williams, J. Holdstock, E. Yague, T. Mulcahy, M. Griffiths, S. Edwards and P. A. Ioannou (2003). "Transgenes encompassing dual-promoter CpG islands from the human< i> TBP</i> and< i> HNRPA2B1</i> loci are resistant to heterochromatin-mediated silencing." Genomics **82**(3): 269-279.

Antoniou, M. N., K. A. Skipper and O. Anakok (2013). "Optimizing retroviral gene expression for effective therapies." Human Gene Therapy **24**(4): 363-374.

Bainbridge, J., M. Tan and R. Ali (2006). "Gene therapy progress and prospects: the eye." Gene Therapy **13**(16): 1191-1197.

Bandaranayake, A. D., C. Correnti, B. Y. Ryu, M. Brault, R. K. Strong and D. J. Rawlings (2011). "Daedalus: a robust, turnkey platform for rapid production of decigram quantities of active recombinant proteins in human cell lines using novel lentiviral vectors." Nucleic Acids Research **39**(21): e143-e143.

Benton, T., T. Chen, M. McEntee, B. Fox, D. King, R. Crombie, T. C. Thomas and C. Bebbington (2002). "The use of UCOE vectors in combination with a preadapted serum free, suspension cell line allows for rapid production of large quantities of protein." Cytotechnology **38**(1-3): 43-46.

Benton, T., T. Chen, M. McEntee, B. Fox, D. King, R. Crombie, T. C. Thomas and C. Bebbington (2002). "The use of UCOE vectors in combination with a preadapted serum free, suspension cell line allows for rapid production of large quantities of protein." Cytotechnology **38**(1): 43-46.

Biffi, A., E. Montini, L. Lorioli, M. Cesani, F. Fumagalli, T. Plati, C. Baldoli, S. Martino, A. Calabria and S. Canale (2013). "Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy." Science **341**(6148): 1233158.

Bosticardo, M., F. Ferrua, M. Cavazzana and A. Aiuti (2014). "Gene therapy for Wiskott-Aldrich Syndrome." Current Gene Therapy **14**(6): 413-421.

Bouard, D., N. Alazard-Dany and F. L. Cosset (2009). "Viral vectors: from virology to transgene expression." British Journal of Pharmacology **157**(2): 153-165.

Brendel, C., U. Müller-Kuller, S. Schultze-Strasser, S. Stein, L. Chen-Wichmann, A. Krattenmacher, H. Kunkel, A. Dillmann, M. Antoniou and M. Grez (2011). "Physiological regulation of transgene expression by a lentiviral vector containing the A2UCOE linked to a myeloid promoter." Gene Therapy **19**(10): 1018-1029.

Brendel, C., U. Müller-Kuller, S. Schultze-Strasser, S. Stein, L. Chen-Wichmann, A. Krattenmacher, H. Kunkel, A. Dillmann, M. Antoniou and M. Grez (2012). "Physiological regulation of transgene expression by a lentiviral vector containing the A2UCOE linked to a myeloid promoter." Gene Therapy **19**(10): 1018-1029.

Cartier, N., S. Hacein-Bey-Abina, C. C. Bartholomae, G. Veres, M. Schmidt, I. Kutschera, M. Vidaud, U. Abel, L. Dal-Cortivo and L. Caccavelli (2009). "Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy." Science **326**(5954): 818-823.

Cavazza, A., A. Moiani and F. Mavilio (2013). "Mechanisms of retroviral integration and mutagenesis." Human Gene Therapy **24**(2): 119-131.

Cavazzana-Calvo, M., A. Fischer, S. Hacein-Bey-Abina and A. Aiuti (2012). "Gene therapy for primary immunodeficiencies: Part 1." Current Opinion in Immunology **24**(5): 580-584.

Cavazzana-Calvo, M., S. Hacein-Bey, G. de Saint Basile, F. Gross, E. Yvon, P. Nusbaum, F. Selz, C. Hue, S. Certain and J.-L. Casanova (2000). "Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease." Science **288**(5466): 669-672.

Charneau, P., M. Alizon and F. Clavel (1992). "A second origin of DNA plus-strand synthesis is required for optimal human immunodeficiency virus replication." Journal of Virology **66**(5): 2814-2820.

Coil, D. A. and A. D. Miller (2004). "Phosphatidylserine is not the cell surface receptor for vesicular stomatitis virus." Journal of Virology **78**(20): 10920-10926.

Delenda, C. (2004). "Lentiviral vectors: optimization of packaging, transduction and gene expression." Journal of Gene Medicine **6**(S1): S125-S138.

Demaison, C., K. Parsley, G. Brouns, M. Scherr, K. Battmer, C. Kinnon, M. Grez and A. J. Thrasher (2002). "High-level transduction and gene expression in hematopoietic repopulating cells using a human imunodeficiency virus type 1-based lentiviral vector containing an internal spleen focus forming virus promoter." Human Gene Therapy **13**(7): 803-813.

Ellis, J. (2005). "Silencing and variegation of gammaretrovirus and lentivirus vectors." Human Gene Therapy **16**(11): 1241-1246.

Escors, D. and K. Breckpot (2010). "Lentiviral vectors in gene therapy: their current status and future potential." Archivum Immunologiae et Therapiae Experimentalis **58**(2): 107-119.

Fuks, F. (2005). "DNA methylation and histone modifications: teaming up to silence genes." Current Opinion in Genetics & Development **15**(5): 490-495.

Furger, A., J. Monks and N. J. Proudfoot (2001). "The retroviruses human immunodeficiency virus type 1 and Moloney murine leukemia virus adopt radically different strategies to regulate promoter-proximal polyadenylation." Journal of Virology **75**(23): 11735-11746.

Gascón, S., J. A. Paez-Gomez, M. Díaz-Guerra, P. Scheiffele and F. G. Scholl (2008). "Dual-promoter lentiviral vectors for constitutive and regulated gene expression in neurons." Journal of Neuroscience Methods **168**(1): 104-112.

Gaspar, H. B., K. L. Parsley, S. Howe, D. King, K. C. Gilmour, J. Sinclair, G. Brouns, M. Schmidt, C. Von Kalle and T. Barington (2004). "Gene therapy of X-linked severe combined immunodeficiency by use of a pseudotyped gammaretroviral vector." The Lancet **364**(9452): 2181-2187.

Greenberg, K. P., S. F. Geller, D. V. Schaffer and J. G. Flannery (2007). "Targeted transgene expression in Müller glia of normal and diseased retinas using lentiviral vectors." Investigative Iphthalmology & Visual Science **48**(4): 1844.

Griesenbach, U. and E. W. Alton (2013). "Moving forward: cystic fibrosis gene therapy." Human Molecular Genetics **22**(R1): R52-R58.

Hacein-Bey-Abina, S., F. Le Deist, F. Carlier, C. Bouneaud, C. Hue, J.-P. De Villartay, A. J. Thrasher, N. Wulffraat, R. Sorensen and S. Dupuis-Girod (2002). "Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy." New England Journal of Medicine **346**(16): 1185-1193.

Hacein-Bey-Abina, S., C. Von Kalle, M. Schmidt, M. McCormack, N. Wulffraat, P. Leboulch, A. Lim, C. Osborne, R. Pawliuk and E. Morillon (2003). "LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1." Science **302**(5644): 415-419.

Howe, S. J., M. R. Mansour, K. Schwarzwaelder, C. Bartholomae, M. Hubank, H. Kempski, M. H. Brugman, K. Pike-Overzet, S. J. Chatters and D. de Ridder (2008). "Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients." Journal of Clinical Investigation **118**(9): 3143.

Jones-Villeneuve, E., M. W. McBURNEY, K. A. Rogers and V. I. Kalnins (1982). "Retinoic acid induces embryonal carcinoma cells to differentiate into neurons and glial cells." Journal of Cell Biology **94**(2): 253-262.

Kappes, J. C. and X. Wu (2001). "Safety considerations in vector development." Somatic Cell and Molecular Genetics **26**(1-6): 147-158.

Katz, R. A. and A. M. Skalka (1994). "The retroviral enzymes." Annual Review of Biochemistry **63**(1): 133-173.

Knight, S., F. Zhang, U. Mueller-Kuller, M. Bokhoven, A. Gupta, T. Broughton, S. Sha, M. N. Antoniou, C. Brendel and M. Grez (2012). "Safer, silencing-resistant lentiviral vectors: optimization of the ubiquitous chromatin-opening element through elimination of aberrant splicing." Journal of Virology **86**(17): 9088-9095.

Lai, Z. and R. O. Brady (2002). "Gene transfer into the central nervous system in vivo using a recombinanat lentivirus vector." Journal of Neuroscience Research **67**(3): 363-371.

Laker, C., J. Meyer, A. Schopen, J. Friel, C. Heberlein, W. Ostertag and C. Stocking (1998). "Host cis-mediated extinction of a retrovirus permissive for expression in embryonal stem cells during differentiation." Journal of Virology **72**(1): 339-348.

Lienert, F., C. Wirbelauer, I. Som, A. Dean, F. Mohn and D. Schübeler (2011). "Identification of genetic elements that autonomously determine DNA methylation states." Nature Genetics **43**(11): 1091-1097.

Loeb, J. E., W. S. Cordier, M. E. Harris, M. D. Weitzman and T. J. Hope (1999). "Enhanced expression of transgenes from adeno-associated virus vectors with the woodchuck hepatitis virus posttranscriptional regulatory element: implications for gene therapy." Human Gene Therapy **10**(14): 2295-2305.

Maguire, A. M., F. Simonelli, E. A. Pierce, E. N. Pugh Jr, F. Mingozzi, J. Bennicelli, S. Banfi, K. A. Marshall, F. Testa and E. M. Surace (2008). "Safety and efficacy of gene transfer for Leber's congenital amaurosis." New England Journal of Medicine **358**(21): 2240-2248.

Miyoshi, H., U. Blömer, M. Takahashi, F. H. Gage and I. M. Verma (1998). "Development of a self-inactivating lentivirus vector." Journal of Virology **72**(10): 8150-8157.

Müller-Kuller, U., M. Ackermann, S. Kolodziej, C. Brendel, J. Fritsch, N. Lachmann, H. Kunkel, J. Lausen, A. Schambach and T. Moritz (2015). "A minimal ubiquitous chromatin opening element (UCOE) effectively prevents silencing of juxtaposed heterologous promoters by epigenetic remodeling in multipotent and pluripotent stem cells." Nucleic Acids Research: gkv019.

Naldini, L., U. Blömer, P. Gallay, D. Ory, R. Mulligan, F. H. Gage, I. M. Verma and D. Trono (1996). "In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector." Science **272**(5259): 263-267.

Ott, M. G., M. Schmidt, K. Schwarzwaelder, S. Stein, U. Siler, U. Koehl, H. Glimm, K. Kühlcke, A. Schilz and H. Kunkel (2006). "Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1." Nature Medicine **12**(4): 401-409.

Ott, M. G., R. Seger, S. Stein, U. Siler, D. Hoelzer and M. Grez (2007). "Advances in the treatment of chronic granulomatous disease by gene therapy." Current Gene Therapy **7**(3): 155-161.

Pfaff, N., N. Lachmann, M. Ackermann, S. Kohlscheen, C. Brendel, T. Maetzig, H. Niemann, M. N. Antoniou, M. Grez and A. Schambach (2013). "A ubiquitous chromatin opening element prevents transgene silencing in pluripotent stem cells and their differentiated progeny." Stem Cells **31**(3): 488-499.

Pluta, K. and M. Kacprzak (2009). "Use of HIV as a gene transfer vector." Acta Biochimica Polonica **56**(4): 531-595.

Poccia, F., L. Battistini, B. Cipriani, G. Mancino, F. Martini, M. L. Gougeon and V. Colizzi (1999). "Phosphoantigen-reactive Vγ9Vδ2 T lymphocytes suppress in vitro human immunodeficiency virus type 1 replication by cell-released antiviral factors including CC chemokines." Journal of Infectious Diseases **180**(3): 858-861.

Rattray, A. J. and J. J. Champoux (1989). "Plus-strand priming by Moloney murine leukemia virus: the sequence features important for cleavage by RNase H." Journal of Molecular Biology **208**(3): 445-456.

Razin, A. (1998). "CpG methylation, chromatin structure and gene silencing—a three-way connection." EMBO Journal **17**(17): 4905-4908.

She, X., C. A. Rohl, J. C. Castle, A. V. Kulkarni, J. M. Johnson and R. Chen (2009). "Definition, conservation and epigenetics of housekeeping and tissue-enriched genes." BMC Genomics **10**(1): 269.

Simonelli, F., A. M. Maguire, F. Testa, E. A. Pierce, F. Mingozzi, J. L. Bennicelli, S. Rossi, K. Marshall, S. Banfi and E. M. Surace (2009). "Gene therapy for Leber's congenital amaurosis is safe and effective through 1.5 years after vector administration." Molecular Therapy **18**(3): 643-650.

Speers, W. C., J. W. Gautsch and F. J. Dixon (1980). "Silent infection of murine embryonal carcinoma cells by Moloney murine leukemia virus." Virology **105**(1): 241-244.

Stein, S., M. G. Ott, S. Schultze-Strasser, A. Jauch, B. Burwinkel, A. Kinner, M. Schmidt, A. Krämer, J. Schwäble and H. Glimm (2010). "Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease." Nature Medicine **16**(2): 198-204.

Talbot, G. E., S. N. Waddington, O. Bales, R. C. Tchen and M. N. Antoniou (2009). "Desmin-regulated lentiviral vectors for skeletal muscle gene transfer." Molecular Therapy **18**(3): 601-608.

Teich, N. M., R. A. Weiss, G. R. Martin and D. R. Lowy (1977). "Virus-Infection of Murine Teratocarcinoma Stem-Cell Lines." Cell **12**(4): 973-982.

Teich, N. M., R. A. Weiss, G. R. Martin and D. R. Lowy (1977). "Virus infection of murine teratocarcinoma stem cell lines." Cell **12**(4): 973-982.

Thomson, J. P., P. J. Skene, J. Selfridge, T. Clouaire, J. Guy, S. Webb, A. R. Kerr, A. Deaton, R. Andrews and K. D. James (2010). "CpG islands influence chromatin structure via the CpG-binding protein Cfp1." Nature **464**(7291): 1082-1086.

Verma, I. M. and N. Somia (1997). "Gene therapy-promises, problems and prospects." Nature **389**(6648): 239-242.

Vogt, V. M. and M. N. Simon (1999). "Mass determination of Rous sarcoma virus virions by scanning transmission electron microscopy." Journal of Virology **73**(8): 7050-7055.

Watanabe, S. and H. M. Temin (1982). "Encapsidation sequences for spleen necrosis virus, an avian retrovirus, are between the 5'long terminal repeat and the start of the gag gene." Proceedings of the National Academy of Sciences **79**(19): 5986-5990.

Williams, S., T. Mustoe, T. Mulcahy, M. Griffiths, D. Simpson, M. Antoniou, A. Irvine, A. Mountain and R. Crombie (2005). "CpG-island fragments from the HNRPA2B1/CBX3 genomic locus reduce silencing and enhance transgene expression from the hCMV promoter/enhancer in mammalian cells." BMC Biotechnology **5**(1): 17.

Williams, S., T. Mustoe, T. Mulcahy, M. Griffiths, D. Simpson, M. Antoniou, A. Irvine, A. Mountain and R. Crombie (2005). "CpG-island fragments from the HNRPA 2 B 1/CBX 3 genomic locus reduce silencing and enhance transgene expression from the hCMV promoter/enhancer in mammalian cells." BMC Biotechnology **5**(1): 17.

Yao, S., T. Sukonnik, T. Kean, R. R. Bharadwaj, P. Pasceri and J. Ellis (2004). "Retrovirus silencing, variegation, extinction, and memory are controlled by a dynamic interplay of multiple epigenetic modifications." Molecular Therapy **10**(1): 27-36.

Yee, T., K. Pasi, P. Lilley and C. Lee (1999). "Factor VIII inhibitors in haemophiliacs: a single-centre experience over 34 years, 1964–97." British Journal of Haematology **104**(4): 909-914.

Yoon, J.-S., M.-Y. Lee, J.-S. Lee, C. S. Park, H.-J. Youn and J.-H. Lee (2009). "Bis is involved in glial differentiation of p19 cells induced by retinoic Acid." Korean Journal of Physiology & Pharmacology **13**(3): 251-256.

Yu, J.-C., M. A. Nash, C. Santiago and W. F. Marzluff (1986). "Structure and expression of a second sea urchin U1 RNA gene repeat." Nucleic Acids Research **14**(24): 9977-9988.

Zaiss, A.-K., S. Son and L.-J. Chang (2002). "RNA 3′ readthrough of oncoretrovirus and lentivirus: implications for vector safety and efficacy." Journal of Virology **76**(14): 7209-7219.

Zhang, F., A. R. Frost, M. P. Blundell, O. Bales, M. N. Antoniou and A. J. Thrasher (2010). "A ubiquitous chromatin opening element (UCOE) confers resistance to DNA methylation–mediated silencing of lentiviral vectors." Molecular Therapy **18**(9): 1640-1649.

Zhang, F., S. I. Thornhill, S. J. Howe, M. Ulaganathan, A. Schambach, J. Sinclair, C. Kinnon, H. B. Gaspar, M. Antoniou and A. J. Thrasher (2007). "Lentiviral vectors containing an enhancer-less ubiquitously acting chromatin opening element (UCOE) provide highly reproducible and stable transgene expression in hematopoietic cells." Blood **110**(5): 1448-1457.

Zufferey, R., J. E. Donello, D. Trono and T. J. Hope (1999). "Woodchuck hepatitis virus posttranscriptional regulatory element enhances expression of transgenes delivered by retroviral vectors." Journal of Virology **73**(4): 2886-2892.

Zufferey, R., T. Dull, R. J. Mandel, A. Bukovsky, D. Quiroz, L. Naldini and D. Trono (1998). "Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery." Journal of Virology **72**(12): 9873-9880.

http://envmedical.com

http://hivbook.com/tag/viral-genome/